



UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE INFORMÁTICA

Moodvie

Recomendador individual para Facebook de películas basado en
temperamentos y estados de ánimo

Autores:

Aleix Garrido Oberink
Juan Francisco Peña Suárez
Sergio Salmerón Majadas

Director: Héctor Gómez Gauchía
Curso académico: 2010/2011
Proyecto de Sistemas Informáticos

AUTORIZACIONES

Autorizamos a la Universidad Complutense de Madrid a utilizar y/o difundir con fines académicos y no comerciales, siempre mencionando expresamente a sus autores, tanto la propia memoria como el código, la documentación y/o el prototipo desarrollado.

PALABRAS CLAVE

Sistemas de recomendación, Computación Centrada en el Usuario (HCC), Computación Afectiva, Razonamiento basado en casos (CBR), Redes Sociales, Aplicación Facebook, web crawler.

RESUMEN

Los sistemas de recomendación son cada vez más necesarios para ayudar a los usuarios a filtrar la enorme cantidad de información de la que se dispone. En el sector cinematográfico, donde cada semana varía la cartelera de películas proyectadas, pueden resultar incluso más útiles, puesto que para muchos aficionados al cine es complicado estar al tanto sobre cada uno de los filmes que se estrenen.

Por otro lado, el auge de las redes sociales en general, y de Facebook en particular, otorga un público potencial que de otro modo sería difícilmente alcanzable, aprovechando así su capacidad viral de propagación.

Ante estas dos razones surge la idea de desarrollar una aplicación para Facebook que realice recomendaciones sobre películas que se proyecten en la actualidad en alguna sala de cine cercana, bien ajustándose a los gustos del usuario, o bien priorizando el estado de ánimo que el usuario desea tener tras ver el filme.

SUMMARY

Recommendation systems have become increasingly necessary to help out users filtering the enormous amount of information available. In the movies sector, where movie listings change every week, they can even be more useful, since it might be difficult for many movie enthusiasts to be up to date in relation with all the movies being released.

On the other hand, the success of social networks in general, and of Facebook in particular, provides a potential amount of public that would be otherwise hardly reachable, taking advantage of the viral propagation capabilities inherent to such.

Based on these two concepts, we have developed a Facebook application that can recommend films played on local cinemas, either selecting a movie depending of user likings, or prioritizing the mood the users want to have after watching the movie.

Tabla de contenidos

1. INTRODUCCIÓN.....	11
1.1. ¿QUÉ PELÍCULA DEBERÍA IR A VER?.....	11
1.2. CONTEXTO Y NECESIDADES.....	11
1.3. FUNCIONES GENERALES.....	12
2. TEMPERAMENTOS.....	15
2.1. ARTESANO	15
2.2. GUARDIÁN	15
2.3. IDEALISTA.....	15
2.4. RACIONAL	15
3. CONOCIMIENTO	17
3.1. INTRODUCCIÓN Y ADQUISICIÓN DEL CONOCIMIENTO	17
3.2. TABLA DE PELÍCULAS.....	17
3.3. TABLA DE TEMPERAMENTOS	18
3.3.1. <i>Tabla 1. Temperamento/Género en Cines</i>	19
3.3.2. <i>Tabla 2. Temperamento/Género por gustos</i>	20
3.3.3. <i>Tabla 3. Tabla Temperamento/Género combinada.</i>	21
4. HERRAMIENTA DE EXTRACCIÓN.....	23
4.1. EXTRACCIÓN DE INFORMACIÓN.....	23
4.1.1. <i>Introducción</i>	23
4.1.2. <i>Información deseada</i>	23
4.1.3. <i>Búsqueda de fuentes</i>	24
4.1.4. <i>Datos obtenidos</i>	24
4.1.5. <i>Elaboración del extractor</i>	24
4.1.6. <i>Funcionamiento</i>	25
4.1.7. <i>TOR</i>	36
4.2. CATEGORIZADOR DE GÉNEROS	37
4.2.1. <i>Ejemplo de resultados</i>	40
4.2.2. <i>Problemas y errores. Soluciones</i>	40
5. SISTEMA RECOMENDADOR.....	42
5.1. RECOMENDADOR TEMPERAMENTO/GÉNERO O RECOMENDADORMIX.	43
5.2. RECOMENDADOR GÉNERO/GÉNERO GLOBAL	43
5.3. RECOMENDADOR TEMPERAMENTO/TEMPERAMENTO.....	45
6. APLICACIÓN WEB PARA FACEBOOK	47
6.1. INTRODUCCIÓN	47
6.2. REGISTRO EN FACEBOOK	47
6.3. CASOS DE USO CONTEMPLADOS EN LA APLICACIÓN WEB	48
6.3.1. <i>Registrarse en la aplicación</i>	48
6.3.2. <i>Responder el test</i>	48
6.3.3. <i>Introducir la provincia</i>	48
6.3.4. <i>Ver las películas en cartelera</i>	48
6.3.5. <i>Ver información sobre los cines</i>	49
6.3.6. <i>Conocer cuándo se proyecta cada película</i>	49
6.3.7. <i>Obtener una recomendación</i>	49
6.4. DESARROLLO DE LA APLICACIÓN WEB	49
6.4.1. <i>Estructura</i>	49
6.4.2. <i>Descripción de los archivos</i>	50
6.4.3. <i>Problemas encontrados</i>	51

7. EXPERIMENTOS Y RESULTADOS	52
7.1. LISTA DE PELÍCULAS	52
7.2. LISTA IDEAL DE RECOMENDACIÓN	52
7.2.1. <i>Resultados de los recomendadores</i>	53
7.2.2. <i>Observaciones</i>	58
8. RECOMENDACIÓN UNIFICADA	59
8.1. PROCESO DE UNIFICACIÓN	59
8.1.2. <i>Pesos de unificación individualizados</i>	61
8.1.3. <i>Recomendación unificada individualizada</i>	62
9. RECOMENDADOR CON ESTADOS DE ÁNIMO	63
9.1. IDEAS PROPUESTAS	63
9.1.1. <i>Proceso de recomendación con estado de ánimo objetivo</i>	63
10. DIAGRAMAS DEL RECOMENDADOR	66
10.1.1. <i>Diagrama de clases</i>	66
10.1.2. <i>Diagrama de secuencia</i>	67
11. FEEDBACK DE LOS USUARIOS	69
11.1. FORMULARIO	69
11.2. RESULTADOS OBTENIDOS	70
11.3. SUGERENCIAS	71
APENDICE. TABLA DE TEMPERAMENTOS	73
ANEXO I. DOCUMENTACIÓN DEL SISTEMA IMPLEMENTADO	75
I.1. MANUAL DEL USUARIO	75
I.1.1. <i>Usuario común</i>	75
I.1.2. <i>Administrador</i>	81
I.2. MANUAL DEL SISTEMA	83
I.2.1. <i>Requisitos del sistema</i>	83
I.2.2. <i>Instrucciones de instalación</i>	83
ANEXO II. DISEÑO DE LA BASE DE DATOS	86
II.1. MODELO ENTIDAD RELACIÓN	86
II.2. MODELO RELACIONAL	87
II.2.1. <i>Índices</i>	89
II.2.2. <i>Foreign Keys</i>	90

1 INTRODUCCIÓN

1.1 *¿Qué película debería ir a ver?*

El propósito del desarrollo de este proyecto es dar solución a esta pregunta desde el enfoque de la computación afectiva y centrada en humanos. La mayoría de sistemas de recomendación individual en el ámbito del cine están sometidos al problema del arranque en frío, donde el usuario es requerido de cierto conocimiento previo, así como de una dedicación de tiempo a la configuración de sus gustos y con la dependencia del filtrado colaborativo. Esto dificulta la recomendación para usuarios esporádicos o casuales, que no tienen un amplio conocimiento del dominio como para responder a preguntas sobre sus gustos.

Para tratar de solventar estos problemas, se utiliza la recomendación basada en temperamentos y estados de ánimo. Con un único usuario y un conjunto de películas resulta suficiente para realizar una recomendación. Del usuario se extrae su temperamento en base a un pequeño test realizado, y de las películas se obtienen sus géneros mediante el uso de CBR-textual. Es con esta información, analizada y tratada correctamente, con la que se realizan las recomendaciones, evitando así el filtrado colaborativo y el arranque en frío antes mencionados.

El conjunto de películas sobre los que realiza la recomendación es el de aquellas que se estén proyectando, actualmente, en algún cine de la provincia del usuario. Para la actualización de la base de datos que contiene dichas películas se utiliza un web crawler que hemos desarrollado específicamente para dicho propósito, y que ha sido utilizado por otro proyecto de SSII, también de recomendación de películas, pero con criterios diferentes. Esto ha acarreado un trabajo extraordinario de coordinación con el otro grupo, para adaptar la información extraída a las necesidades de ambos proyectos.

1.2 *Contexto y necesidades*

En la sociedad de la información, donde la tecnología a día de hoy -y cada vez más- permite el acceso a un volumen de datos creciente y de rápida difusión, el usuario puede verse desbordado por la cantidad de información disponible, o por la tarea de discernir entre la información que merece la pena y la que no. Por eso mismo, un recomendador puede ser de gran utilidad, ya que facilita la tarea al usuario que desea algo pero se pierde en el intento. El recomendador le ofrece la información “masticada” y orientada exclusivamente hacia la persona que está utilizándolo.

Por otro lado, el creciente uso de las redes sociales nos permite explotar una brecha creada en estos últimos años, pudiendo acercar nuestro proyecto al público general de una manera cómoda y sencilla. Facebook alcanza en Mayo de 2011 los 600 millones de usuarios, de los cuales contamos aproximadamente con 14 millones de usuarios en España. En cuanto al ámbito del cine, las taquillas españolas vendieron un total de 97,2 millones de entradas en 2010. Analizando el beneficio que genera una industria como la del cine y observando el volumen creciente de usuarios de Facebook, podemos concluir que una aplicación de recomendación de cine en Facebook puede cubrir las necesidades de mucha gente.

Si tenemos en cuenta, además, la querencia general de la población hacia la obtención de mejores resultados con menos esfuerzo, es lógico pensar que una aplicación en la que no se requiere conocimiento externo por parte del usuario y que ofrece buenos resultados, puede ser idónea. Una persona que no haya visto ni una película en su vida, puede usar la aplicación del mismo modo que un consolidado cinéfilo, no hay distinción. Con otro tipo de aplicaciones, una persona sin conocimiento en la materia solamente obtendría malos resultados en la recomendación. En el polo opuesto, a una persona con extremo conocimiento en la materia, le resultaría tedioso tener que configurar demasiados apartados en una aplicación para obtener la recomendación.

En cambio, con el uso del sistema de recomendación basado en temperamentos que se desarrolla en este proyecto, el usuario no necesita conocimiento externo; simplemente el conocimiento de uno mismo es suficiente para rellenar el test inicial y poder acceder a las recomendaciones.

En conclusión, si agrupamos las ideas expuestas y las combinamos, nos encontramos con la meta de este proyecto. Un recomendador eficaz en sus resultados, fácil de usar y accesible para mucha gente gracias a las redes sociales. Todo ello elaborado desde el enfoque de la computación afectiva. A lo largo de la memoria se detallan los múltiples experimentos realizados hasta alcanzar la funcionalidad deseada, con los fallos y aciertos de cada uno de los enfoques.

1.3 Funciones generales

Como idea introductoria de las funciones generales que se plantean en el proyecto podemos mencionar los siguientes puntos.

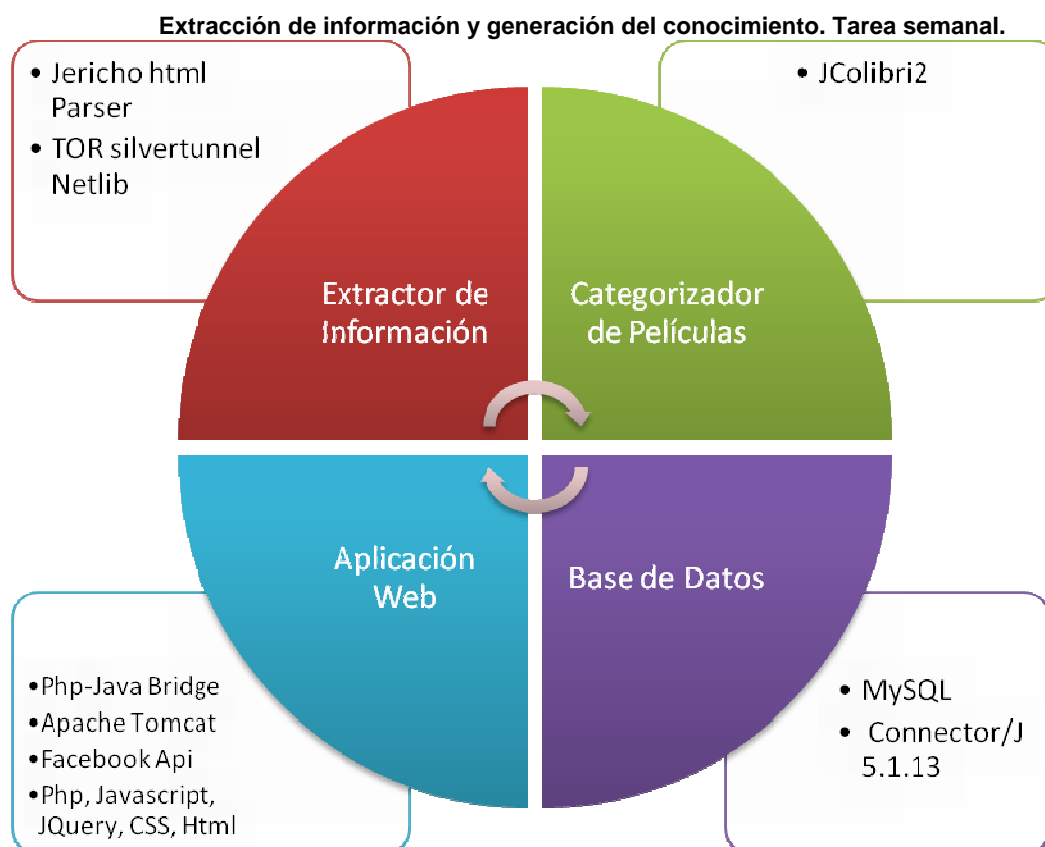
Primeramente, para realizar un recomendador, es necesario información y conocimiento. Con la intención de obtener películas y detalles sobre ellas, se procesa, con la ayuda de una araña web, la página de “La Guía del Ocio” y se procede al análisis sintáctico de su contenido. De este modo se obtiene, en un proceso bastante cercano a la automatización, información sobre cada una de las películas que se estén proyectando en ese momento en salas de cine españolas, así como información sobre dichas salas –y los pases de cada una de ellas. El desarrollo de la araña es vital, puesto que al tratarse de películas en cartelera, la información es cambiante con periodicidad semanal, y su procesamiento manual sería enormemente costoso e ineficiente.

Una vez extraída la información, es necesario dotar de determinada semántica a todos los datos obtenidos. Es entonces cuando se realiza un procesamiento CBR textual de la sinopsis de los filmes obtenidos en el paso anterior. Con ello se realizan una serie de categorizaciones, transformando pues la información en conocimiento, que posteriormente servirá para realizar las recomendaciones deseadas.

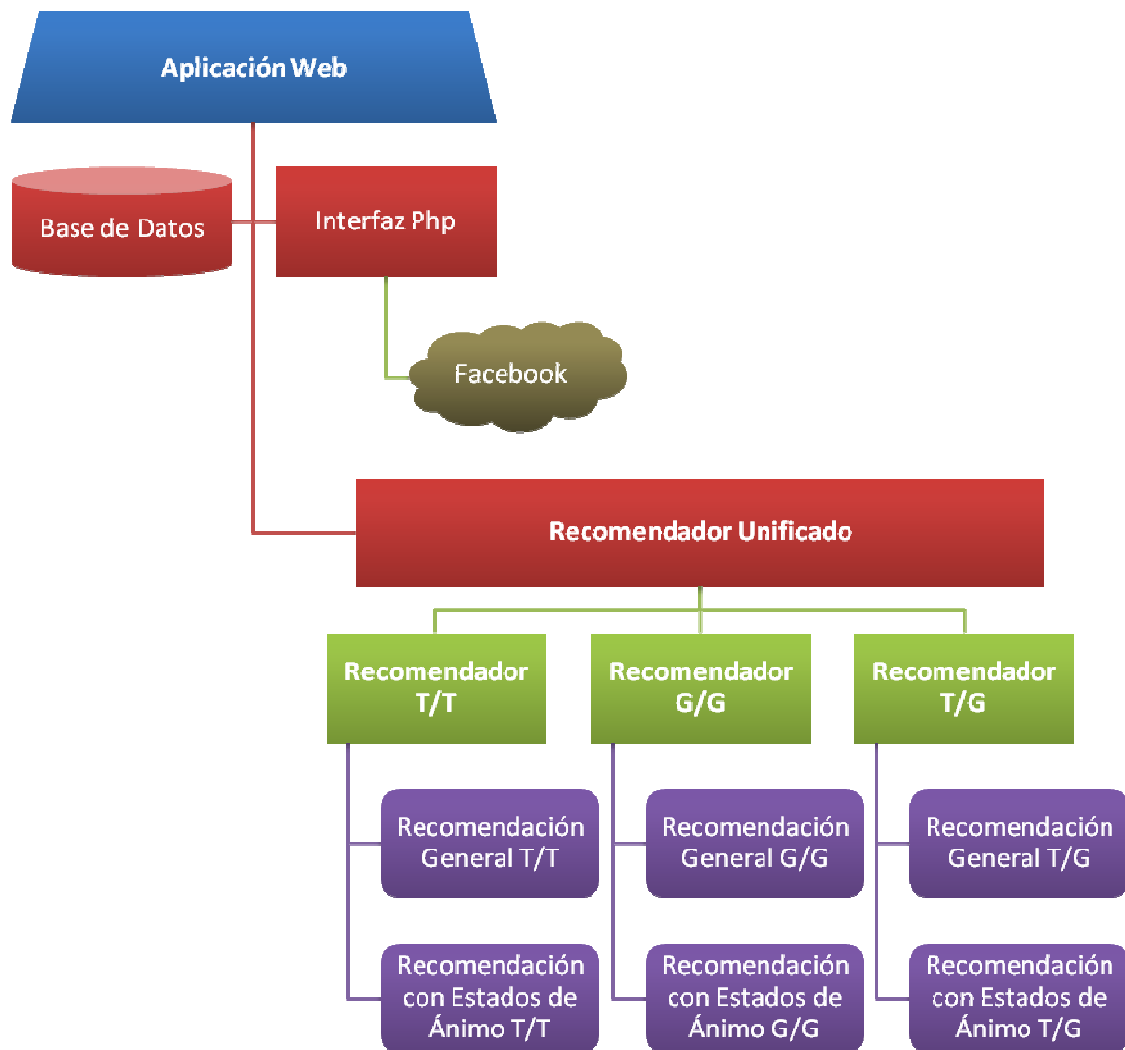
Es en este momento, cuando el conocimiento objetivo ya se encuentra disponible para la tarea de recomendación, cuando se hace necesario en conocimiento fuente. Para ello, se dota al usuario de un test de 16 preguntas en una interfaz similar a la de Facebook, en la que se le supone acostumbrado. Una vez contestadas estas preguntas, el sistema recomendador genera un conocimiento fuente, que en este caso es el temperamento del usuario. En base a este, y filtrando por la provincia que haya seleccionado, ya se pueden realizar recomendaciones, bien genéricas, bien basadas en estado de ánimo deseable -que puede seleccionarse al solicitar una recomendación- ajustadas a los gustos del usuario de la aplicación.

Como valor añadido para el usuario, éste puede obtener información sobre cualquiera de las salas de cine o películas del sistema, así como conocer cuándo se proyecta cada una de las películas película en un cine a elección del usuario. Esto hace que, en la misma aplicación, el usuario, además de obtener una recomendación sobre qué película pueda ajustarse más a sus gustos, conozca las salas de cine de su provincia donde proyectan dicho filme (incluso los pases de cada sala), ahorrando búsquedas innecesarias o laboriosas.

Así pues, en el siguiente diagrama se puede visualizar de manera somera el funcionamiento del sistema desarrollado.



Esquema general de funcionamiento de la aplicación.



2 TEMPERAMENTOS

El proyecto consiste en un recomendador de películas basado el temperamento del usuario. El temperamento es la manera natural en la que un ser humano interactúa con el entorno, una cualidad genética innata prácticamente invariable (mientras que el estado de ánimo es una propiedad variable).

Existen multitud de estudios sobre los temperamentos sobre los que podríamos basar la parte de conocimiento de nuestro recomendador. Trabajaremos con las teorías y estudios llevados a cabo por David Keirsey, profesor y psicólogo estadounidense autor de un cuestionario conocido como “KeirseyTemperamentSorter”.

Keirsey divide el temperamento en cuatro temperamentos básicos:

2.1 *Artesano*

La habilidad natural de destacar en algún arte, desde las bellas artes hasta el arte de la dialéctica o los negocios. Amantes de la diversión, realistas pero optimistas, disfrutan del presente. Creativos y espontáneos, en busca de causar impacto en los demás. Confían en sus impulsos, creen que todo les saldrá bien.

2.2 *Guardián*

Piedra angular de la sociedad, personas entregadas a servir y preservar las instituciones. Gran capacidad de gestión de bienes y servicios. Buenos supervisores. Personas leales, obedientes, responsables y muy trabajadoras. Actúan con cautela, buscan seguridad en sus vidas. Adoran la justicia y la pertenencia a un grupo.

2.3 *Idealista*

Pasionales y entregados al crecimiento personal. Se esfuerzan en descubrir su verdadero yo, en mejorar como personas. Disfrutan ayudando a la gente y creando vínculos. Anhelan una vida romántica y la sabiduría. Intensos y espirituales. Gente muy auténtica, querida y ética en sus acciones.

2.4 *Racional*

Resolutores de problemas. Analizan el mundo con intención de entender cómo funcionan las cosas. Escépticos y pragmáticos. Orgullosos de su ingenio. Líderes estratégicos, con fuerza de voluntad. Tranquilos y serenos, se dejan guiar por la lógica. Disfrutan con la tecnología. A veces se ven como personas frías y distantes.

El temperamento de cada persona, o usuario de nuestra aplicación, será representado mediante una combinación única de porcentajes de los cuatro temperamentos básicos.

Por ejemplo: Usuario X = 10% Artesano, 30% Guardián, 25% Idealista y 35% Racional), llevando a cabo cálculos realizados a partir de los resultados que se hayan obtenido tras rellenar el test de Keirsey.

La esencia de cada temperamento está contenida en la tabla de temperamentos anotados. Esta tabla de resumen ha sido una herramienta de consulta habitual a la hora de guiar la parte de conocimiento de nuestro recomendador. Con una lectura breve se pueden extraer las características que definen cada temperamento agilizando la tarea establecer relaciones entre los géneros de películas y los temperamentos. A continuación se muestra un pequeño extracto de la tabla completa, que puede consultarse en el Apéndice I. Tabla de Temperamentos.

Artesano	Guardián	Idealista	Racional
Artísticos	Confiables	Empáticos	Ingeniosos
Audaces	Caritativos	Benévolos	Autónomos
Adaptables	Respetables	Auténticos	Resueltos

3 CONOCIMIENTO

3.1 *Introducción y adquisición del conocimiento*

Para poder recomendar, es necesario tener un conocimiento previo. Como ya se ha comentado en la sección 1.-Introducción, diferenciamos entre conocimiento fuente y conocimiento objetivo.

El primero de ellos será el relacionado al temperamento del usuario, que nos servirá de guía para buscar un resultado ajustado. Para obtener este conocimiento, basta con que el usuario de la aplicación conteste al test que tiene a su disposición en la aplicación web, e introduzca la provincia en la que se encuentra. Desde ese momento, el conocimiento fuente para dicho usuario estará completado.

El segundo, el llamado conocimiento objetivo, es todo aquel relacionado con el objetivo a recomendar, esto es, la película y la sala de cine. Para ello, en base a la información obtenida mediante el Extractor y almacenada en la base de datos, se realizarán diferentes cálculos que nos permitirán la transformación de dicha información a conocimiento. En un primer paso, veremos cómo se obtiene la clasificación por géneros de cada una de las películas. En un segundo, la manera ideada para la transformación entre géneros y temperamentos, con lo que se podrán realizar las recomendaciones.

3.2 *Tabla de películas*

El conocimiento es la piedra angular del proyecto. En base al conocimiento se realizan todos los cálculos que conciernen a la recomendación. La intención es enlazar la teoría de temperamentos de Keirsey con la información de las películas.

Primeramente hay que valorar qué información tenemos sobre las películas que pueda ser útil para el desarrollo de la base de conocimiento.

Tenemos por un lado el director y los actores. Esta información a priori podría parecer interesante, pero en realidad no aporta demasiado al tipo de recomendación que estamos buscando. Tanto director como actores pueden haber trabajado en muchos tipos de películas, de las cuales, algunas pueden gustar al usuario del recomendador mientras que otras pueden desagradarle.

Con el país de origen o la duración de la película puede pasar algo similar. Una persona puede sentirse inclinada a producciones de ciertas nacionalidades o incluso puede tender a preferir películas que se encuentren dentro de un rango de minutos de metraje, pero el número de excepciones puede ser altísimo y no podríamos realizar una recomendación basándonos en estos parámetros. Como mucho podría añadirse un pequeño modificador, cuyo peso sería insignificante y que probablemente añadiría más ruido que mejora. Por lo tanto descartamos estos datos.

Información que podría ser relevante, es la de los personajes de la película. Si se analiza el temperamento de los personajes que forman parte de la trama, dando más peso a los protagonistas, podríamos crear funciones de similitud entre los roles que participan en la acción y los usuarios. Lamentablemente esta información no se encuentra disponible en ningún lado. El único modo de extraer este tipo de datos sería viendo una a una todas las películas que hay en cartelera, y realizar un estudio exhaustivo. Obviamente, no hay tiempo para un estudio de tales características, así que se debe descartar la idea.

Nos queda por tanto la sinopsis y el género. Estos datos sí que proporcionan información clave a la hora de realizar una recomendación. De la sinopsis se puede sacar información de la película (veremos más adelante cómo se trata la sinopsis mediante CBR textual para extraer un conjunto de géneros). Y tiene lógica pensar que una persona puede sentirse atraída hacia determinados géneros, o determinadas combinaciones de géneros. Por tanto, la información que vamos a usar principalmente para el conocimiento de nuestro recomendador es el género de la película.

3.3 *Tabla de temperamentos*

Para la concepción de la base de conocimiento que relaciona géneros con temperamentos el primer paso es extraer las ideas clave de las teorías de Keirsey. Con esta intención, los miembros del grupo leímos detalladamente las características de los cuatro temperamentos principales, para comprender en detalle el comportamiento de personas con esas características y ponernos en el papel de cada uno de ellos. Nos servimos del uso de la Tabla anotada de Temperamentos, que utilizamos de referencia en muchos puntos del proyecto ya que contiene las palabras clave que definen cada temperamento. Esta tabla puede consultarse en el Apéndice I. Tabla de temperamentos.

Por otro lado, tenemos que comprender en qué consiste cada género de los existentes. Esta tarea es algo más sencilla, ya que los géneros son algo que todos conocemos por películas que ya hemos visto (únicamente una breve lectura a modo de recordatorio es suficiente para repasar los puntos clave de cada uno de ellos). Los géneros que vamos a tener en cuenta como géneros base, y de los que nos vamos a servir a la hora de realizar la recomendación son los siguientes:

- Acción
- Animación
- Aventuras
- Comedia
- Documental
- Drama
- Fantasía
- Romántica
- Terror
- Thriller
- Ciencia Ficción

Si relacionamos el conocimiento extraído de los géneros con el perteneciente a los temperamentos mediante una tabla de conversión, podemos distribuir una serie de pesos que sirvan para la recomendación.

Para la obtención de estos pesos, a partir de los cuales se realizan los cálculos necesarios para el algoritmo recomendador, se han elaborado una serie de tablas, cada una con un propósito, que serán explicadas en los puntos siguientes. Mediante estas tablas de conversión conseguimos transformar el acceso desde géneros a temperamentos y viceversa, con lo que obtenemos una gran potencia en la aplicación.

3.3.1 Tabla 1. Temperamento/Género en Cines

Esta es la primera tabla que forma parte de la estructura y organización del conocimiento. En ella tenemos los géneros en las columnas y los temperamentos por filas. Cada columna de género simula una sala de cine, y en ella distribuimos porcentualmente los temperamentos que se encuentran en la sala (por ejemplo, en una sala en la que se proyecta una película que es 100% de acción, encontraríamos una distribución de personas en las que habría un 45% de Artesanos, 21'67% de Idealistas, 21'67% de Guardianes y 11'67% de Racionales). Por cada género repartimos un 100% entre los cuatro temperamentos, obteniendo la siguiente tabla:

Acción		Animación	Aventuras
Artesano	45,00%	31,67%	40,00%
Guardián	21,67%	13,33%	20,00%
Idealista	21,67%	30,00%	26,67%
Racional	11,67%	25,00%	13,33%
Comedia		Documental	Drama
Artesano	25,00%	11,67%	10,00%
Guardián	25,00%	36,67%	31,67%
Idealista	31,67%	20,00%	35,00%
Racional	18,33%	31,67%	23,33%
Fantasía		Romántica	Terror
Artesano	30,00%	13,33%	31,67%
Guardián	22,67%	28,33%	24,00%
Idealista	30,00%	40,00%	24,33%
Racional	17,33%	18,33%	20,00%
Thriller		Ciencia ficción	
Artesano	13,33%	10,00%	
Guardián	38,33%	21,67%	
Idealista	28,33%	28,33%	
Racional	20,00%	40,00%	

Los porcentajes que encontramos en esta tabla son la media resultante de las valoraciones individuales realizadas por los miembros del grupo de proyecto. Cada uno rellenó una tabla por su cuenta, teniendo presente su objetivo.

En un principio se tenía la intención de usar esta tabla como fuente de conocimiento del recomendador. Pero sometida a una serie de pruebas se pudo observar que no era suficiente, ya que devolvía en ocasiones recomendaciones no deseadas.

Las recomendaciones no deseadas se debían al hecho de considerar todos los géneros con el mismo peso. Imaginemos un género determinado X que no guste a ninguno de los temperamentos: el género odiado. La distribución en la sala de cine para una película del género odiado sería de 25% por cada temperamento. Por otro lado supongamos un género Y, que fuese un género que gusta mucho a todos los temperamentos: el género amado. La distribución en una sala en la que se proyecta una película del género amado sería también del 25% para cada uno de los cuatro temperamentos. Entonces, si apareciesen en la cartelera las dos películas X e Y, el recomendador les otorgaría a las dos la misma puntuación, ya aparentemente serían iguales en cuanto a la distribución de porcentajes. Esto nos hace ver que esta tabla no cubre todos los aspectos que cabría tener en cuenta para realizar una recomendación.

3.3.2 Tabla 2. Temperamento/Género por gustos

Al ver que la primera tabla no era suficiente para realizar una recomendación acertada, nos vemos ante la necesidad de cambiar el enfoque de la distribución del conocimiento.

Con esta intención creamos una nueva tabla, que distribuye ahora los porcentajes de manera horizontal, por gustos. En las filas volvemos a encontrar los temperamentos, y en las columnas los géneros, pero ahora el planteamiento es distinto. En esta ocasión suponemos que cada temperamento se le da la oportunidad de ver una película de cada género y darle una puntuación del 1 al 10. Si sumamos los puntos que ha dado en total un temperamento a todos los géneros, luego podemos sacar el porcentaje de puntos que se ha dado a un género de manera individual. Lo vemos en las tablas siguientes:

Esta primera tabla contiene las puntuaciones otorgadas por un temperamento concreto a un género determinado.

	Acción	Animación	Aventuras
Artesano	8	7,6666667	10
Guardián	6,6666667	6	5,6666667
Idealista	5,3333333	7,3333333	6,6666667
Racional	3	7	6,3333333
	Comedia	Documental	Drama
Artesano	6,3333333	3,3333333	3,6666667
Guardián	7	8	9,3333333
Idealista	8	7,6666667	8
Racional	6,6666667	9,3333333	7,3333333
	Fantasía	Romántica	Terror
Artesano	7,6666667	5	7
Guardián	7	6,6666667	5
Idealista	7	10	2,6666667
Racional	6,6666667	6	4,6666667
	Thriller	Ciencia ficción	
Artesano	6,6666667	7,3333333	
Guardián	7,3333333	4,3333333	
Idealista	4,6666667	4,6666667	
Racional	7,6666667	9,3333333	

Esta segunda tabla tiene los porcentajes de distribución de puntuaciones.

	Acción	Animación	Aventuras
Artesano	11,01%	10,55%	13,76%
Guardián	9,13%	8,22%	7,76%
Idealista	7,41%	10,19%	9,26%
Racional	4,05%	9,46%	8,56%
	Comedia	Documental	Drama
Artesano	8,72%	4,59%	5,05%
Guardián	9,59%	10,96%	12,79%
Idealista	11,11%	10,65%	11,11%
Racional	9,01%	12,61%	9,91%
	Fantasia	Romántica	Terror
Artesano	10,55%	6,88%	9,63%
Guardián	9,59%	9,13%	6,85%
Idealista	9,72%	13,89%	3,70%
Racional	9,01%	8,11%	6,31%
	Thriller	Ciencia ficción	
Artesano	9,17%	10,09%	
Guardián	10,05%	5,94%	
Idealista	6,48%	6,48%	
Racional	10,36%	12,61%	

Ahora, si sumamos todos los valores por cada temperamento, alcanzamos el 100%. Hemos distribuido los gustos y de este modo dejamos a un lado el problema encontrado en la tabla 1, donde una película amada y una odiada podían sacar la misma puntuación.

Para llegar a este resultado, nuevamente los tres miembros de grupo han realizado una tabla cada uno, dando puntuaciones a cada género. De las tres tablas resultantes se ha realizado la media, y sobre ella se ha calculado la distribución de puntuaciones o gustos.

Al someter esta tabla a las pruebas de recomendación de modo individual se obtienen buenos resultados, pero demasiado similares entre ellos. Partimos de la base de que una persona es una distribución de temperamentos, y una película es una distribución de géneros. Con estas premisas, al tener que coger muchos porcentajes de muchas casillas de la tabla, todo tiende a unificarse. Por esto, aparte del hecho de que las puntuaciones son muy similares entre ellas, cabe decir que cualquier error en la distribución de géneros de la película podía ser fatal a la hora de recomendar. Por este motivo, la tabla basada en gustos no es suficiente para realizar una recomendación idónea.

3.3.3 Tabla 3. Tabla Temperamento/Género combinada

Ante la ineficiencia de la tabla 1 para discernir entre géneros amados y aborrecidos, y la excesiva similitud de los resultados obtenidos por la tabla 2, nos vemos ante la necesidad de encontrar un nuevo punto de vista.

Ya que los dos planteamientos de las tablas anteriores no eran malos, parece que puede ser buena idea combinarlos. Si aumentamos las distancias de las puntuaciones de la tabla 2 a partir del peso otorgado en la tabla 1, obtenemos mediante unos cálculos la tabla 3 o tabla combinada.

La tabla combinada tiene la misma estructura que la tabla de gustos, es decir, se distribuyen en ella las puntuaciones horizontalmente. La diferencia radica en que ahora las puntuaciones se van a ver acentuadas por la tabla de cines. Es factible pensar que las puntuaciones otorgadas a los géneros por personas de un determinado temperamento tengan mayor o menor relevancia dependiendo de la cantidad de personas que vayan a verlas al cine. En el fondo son datos relacionados, así que lo que vamos a conseguir es una redistribución de gustos con valores relativamente más distanciados.

Lo vemos en la siguiente tabla:

	Acción	Animación	Aventuras
Artesano	18,78%	12,66%	20,86%
Guardián	7,43%	4,11%	5,83%
Idealista	5,40%	10,28%	8,31%
Racional	2,02%	10,12%	4,88%
	Comedia	Documental	Drama
Artesano	8,26%	2,03%	1,91%
Guardián	9,00%	15,09%	15,20%
Idealista	11,84%	7,17%	13,09%
Racional	7,07%	17,10%	9,90%
	Fantasia	Romántica	Terror
Artesano	12,00%	3,48%	11,56%
Guardián	8,16%	9,72%	6,17%
Idealista	9,82%	18,70%	3,03%
Racional	6,68%	6,36%	5,40%
	Thriller	Ciencia ficción	
Artesano	4,64%	3,82%	
Guardián	14,46%	4,83%	
Idealista	6,18%	6,18%	
Racional	8,87%	21,59%	

Como vemos, los límites entre los gustos son mucho más marcados que en la tabla 2. Para llegar a este resultado se ha realizado la siguiente operación, donde *tablaCombinada* es la tabla mostrada en este apartado, *tablaCines* es la tabla 1 y *tablaGustos* es la tabla de puntuaciones mostrada en el apartado de la tabla 2.

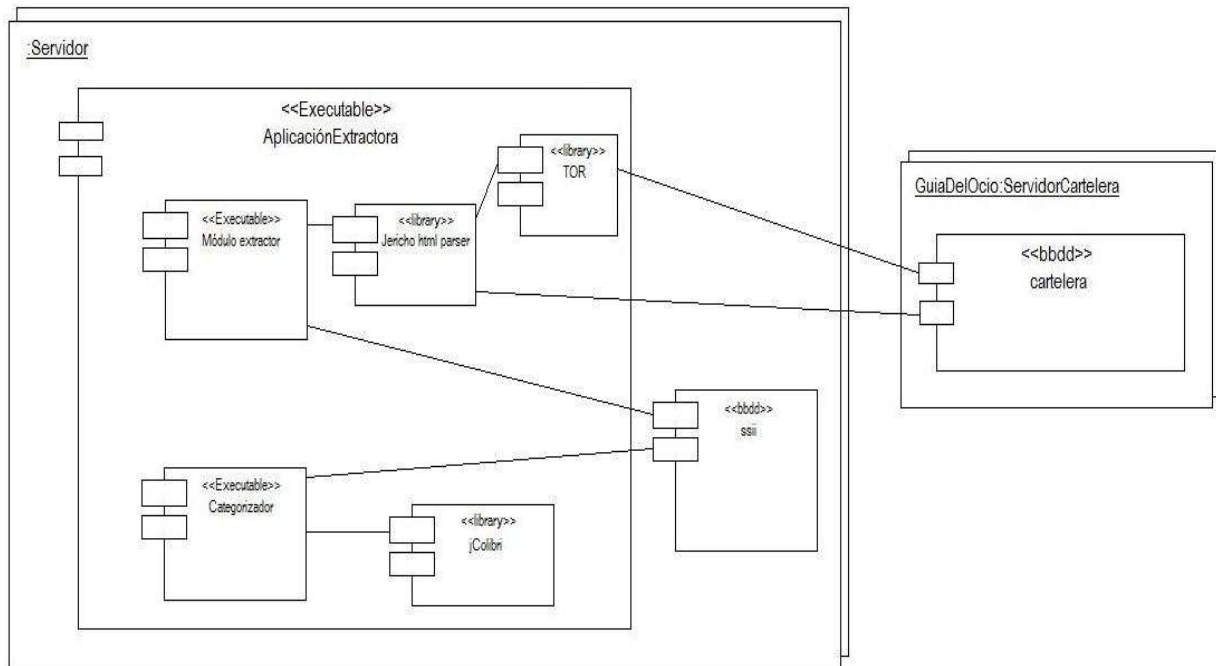
$$tablaCombinada(t, g) = \frac{tablaCines(t, g) * tablaGustos(t, g)}{\sum_{g2=Acción}^{Ciencia Ficción} tablaCines(t, g2) * tablaGustos(t, g2)}$$

Siendo *t* el temperamento y *g* el género de la posición de la tabla.

Al ejecutar las pruebas sobre esta nueva tabla de conocimiento encontramos los mejores resultados obtenidos hasta ahora. La combinación de las dos ideas, cogiendo lo mejor de cada una, solventa los problemas que encontrábamos hasta ahora, y por tanto, pasa a ser la tabla principal usada en nuestro sistema de recomendación y la fuente de conocimiento de la que beben nuestros algoritmos.

4 HERRAMIENTA DE EXTRACCIÓN

Para el proyecto se ha creado una herramienta cuya única funcionalidad es extraer la información y procesarla, dejando la información utilizable por el recomendador. Pueden distinguirse dos módulos dentro de esta aplicación, uno encargado de obtener la información de internet, y otro de su tratamiento. A continuación podemos ver un diagrama explicativo de su estructura y la explicación de cada uno de estos dos módulos:



4.1 Extracción de información

4.1.1 Introducción

Puesto que el proyecto se basa en el manejo de conocimiento, es necesario que cuente con una fuente de información que le proporcione todos los datos necesarios para trabajar. De la calidad de la información extraída depende en gran parte la calidad de los resultados ofrecidos por cualquier sistema basado en conocimiento.

4.1.2 Información deseada

Puesto que nuestro sistema se basa en la teoría de los temperamentos, inicialmente se contempló la posibilidad de realizar la recomendación basándonos principalmente en los personajes de las películas (comparando sus temperamentos con el temperamento del usuario), buscando así la identificación del usuario con el personaje de la película.

Debido a la naturaleza de los proyectos basados en conocimiento, se deberá valorar el uso de cualquier otro tipo de información que pueda reforzar el mecanismo de nuestro proyecto, ya sea para centrar la recomendación sobre él o para refinar el funcionamiento mediante su uso.

Otro factor en tener en cuenta es que la fuente de información nos pueda “garantizar” su permanencia en el tiempo (ya que en cuanto el sistema no cuente con información actualizada, podría ofrecer alguna recomendación errónea) así como películas que estén al alcance de los usuarios (por lo que será necesario contar con una página que ofrezca la cartelera nacional).

4.1.3 Búsqueda de fuentes

La primera fase de la Adquisición del Conocimiento fue la búsqueda de las fuentes. Tras la puesta en común de los sitios encontrados como candidatos a su uso para la extracción de información de ellos, vimos que no había ninguno en el que encontrásemos información relevante sobre los personajes de todas las películas (en algunas, se menciona algo en la sinopsis, pero nunca de una forma estructurada que facilite su extracción para nuestro proyecto).

4.1.4 Datos obtenidos

Una vez estudiadas las propuestas de los miembros, elegimos la página web de La Guía del Ocio (<http://www.guiadelocio.com/>) como principal fuente de información. Puesto que es una empresa que se dedica a la publicación de la oferta del ocio en las ciudades españolas, nos ofrece una garantía de continuidad de servicio, entre el que se encuentra la cartelera nacional.

Otro factor a tener en cuenta a la hora de elegir la página de La Guía del Ocio es la información que ofrece de cada película, ya que nos ofrece una sinopsis bastante amplia de la película, al igual que otra mucha información técnica sobre ella (director, reparto, año, nacionalidad, etc...) e información sobre los cines de cada provincia en los que se emite cada película (ofreciéndonos esto la posibilidad de recomendar a cada usuario sólo películas que se proyecten en su localidad). La obtención de la sinopsis resulta crítica, ya que esta es utilizada para extraer, mediante CBR Textual, una colección de géneros y su porcentaje, complementando así al otorgado por La Guía del Ocio.

4.1.5 Elaboración del extractor

Para implementar el extractor, una de las primeras cosas que se hizo fue buscar material que pudiese resultarnos útil a la hora de manejar muchos archivos HTML. Una de las mejores opciones sería el uso de una librería que nos facilitase y agilizase la navegación por el código de los ficheros que componían la cartelera de La Guía del Ocio. Tras probar algunas, decidimos incluir en el extractor la librería Jericho HTML Parser (<http://jericho.htmlparser.net>), cuyo uso nos permite acceder de forma rápida (y sencilla) a cualquier etiqueta de los archivos HTML (que generalmente son grandes) del portal.

Una vez tenemos una librería que nos facilite la navegación por los HTML, procedemos al estudio de los archivos de cines y películas, viendo cómo podemos extraer la información de ellos. Afortunadamente, el uso de etiquetas exclusivas en los campos en los que se encuentran los datos deseados dentro del código hará que la labor de extracción de dichos datos no se complique en ese aspecto.

Por último (antes de comenzar a implementar el extractor), sabiendo ya la información que podemos extraer, creamos los tipos de datos necesarios para albergar la información obtenida antes de introducirla en la base de datos. Estos tipos de datos serán: Película (que contiene un conjunto de datos de una película), Proyección (que contiene un identificador de película, un día de la semana y una hora) y Cine (que contiene toda la información de un cine al igual que un array de objetos de tipo Proyección).

4.1.6 Funcionamiento

El funcionamiento del extractor se ha dividido en 2 bloques independientes:

4.1.6.1 EXTRACCIÓN E INFORMACIÓN DE PELÍCULAS

Encargado de extraer toda la información de cada película.

4.1.6.1.1 Información extraída por cada película:

- URL de la película dentro de La Guía del Ocio: Elemento que utilizaremos como identificador de las películas durante el proceso de extracción (Una vez terminada la extracción, se generará un número de identificador único para cada película al incluirla en la base de datos).
- Título
- Año
- Duración
- Director
- Intérpretes
- Nacionalidad
- Fecha de estreno
- Género
- Productora
- Guionista
- Distribuidora
- Calificación de edades
- Sinopsis
- Nombre del archivo del cartel descargado

4.1.6.1.2 Funcionamiento detallado

Inicialmente cargamos la página de la cartelera de cada provincia (<http://www.guiadelocio.com/cartelera/<PROVINCIA>?vista=3>), tras cuyo procesamiento almacenamos todos los enlaces a películas contenidos en ella (todas las películas proyectadas en esa provincia). Con todos los enlaces guardados, procedemos a descargar el fichero HTML al que dirige cada uno de ellos y obtener de ahí toda la información de cada película, con la salvedad de realizar, primeramente, una comprobación sobre si dicha película existe ya en la base de datos. En ese caso, el enlace se deshecha y la película no se vuelve a descargar una y otra vez, consiguiendo así una enorme ganancia en tiempo de ejecución, así como en cantidad de tráfico generado.

Página de la cartelera de una provincia:

The screenshot shows the 'Guía del Ocio' website for Madrid. The main navigation bar includes links for 'Ocio en tu zona' and 'Club GDO'. Below this, there are tabs for 'Portada', 'Cine', 'Restaurantes', 'Teatro y danza', 'Conciertos', 'Arte', 'Tarde y noche', 'Niños', 'Ferias', 'Despedidas y fiestas', and 'Cursos y aficiones'. The 'Cine' tab is selected, leading to the 'Cine en Madrid' section. On the left, there is a sidebar with links to 'Estrenos', 'Cartelera', 'Críticas', 'Trailers', 'Próximos estrenos', 'Estrellas', 'Ranking', 'Galerías', 'Personajes', 'Salas', 'Noticias', and 'Archivo películas'. The 'Cartelera' section is active, showing a list of movies with their titles and genres. The list includes 'Agua para elefantes', 'Aguila Roja: La película', 'Alexia', 'Año bisiesto', 'Arthur y la guerra de los mundos', 'Bebés', 'Bestezuelas', 'Bienvenidos al Sur', 'Biutiful', and 'Brutal Box'. On the right, there is a 'Club GDO' section with a sign-up form and a 'Lo + valorado' section listing top-rated movies with their ratings and number of votes. An arrow points from the movie list to the text below.

Enlaces a las páginas con las fichas de las películas

A continuación mostramos un extracto de la estructura de la página web de la cartelera de una provincia, explicando sus diferentes partes y cómo es tratado por el parseador de HTML Jericho.

Extracto de código de la página de la cartelera de una provincia:

```

<div class="gridType06 ftl clear">
  <div class="moduleType101 borderbclr clear">
    <p class="texto withBolo01">
      <strong>
        <a href="/cine/archivo-peliculas/carl-gustav-jung">
          <spam class="titulotexto">Carl Gustav Jung</spam></a>
        </strong>
      </p>
      <br/>
      <p class="texto">
      </p>
    </div>

    <div class="moduleType101 borderbclr clear">
      <p class="texto withBolo01">
        <strong>
          <a href="/cine/archivo-peliculas/delfines-y-ballenas">
            <spam class="titulotexto">Delfines y Ballenas</spam></a>
          </strong>
        </p>
        <br/>
        <p class="texto">
        </p>
      </div>

    <div class="moduleType101 borderbclr clear">
      <p class="texto withBolo01">
        <strong>
          <a href="/cine/archivo-peliculas/dinosaurios-alive">
            <spam class="titulotexto">Dinosaurios Alive</spam></a>
          </strong>
        </p>
        <br/>
        <p class="texto">
        </p>
      </div>

    . . .
  </div>


```

El extracto de código arriba mostrado es un fragmento de la parte de código que nos interesa de la página de la cartelera de una provincia. El divisor con el identificador “gridType06 ftlclear” contiene los enlaces a todas las películas de las provincias (cada uno dentro de un divisor de clase “moduleType101 borderbclrclear”). Solo debemos acceder a ellos por medio del uso de la librería y guardar la dirección del enlace de cada película.

Página de la ficha de una película:

Chico y Rita Archivo

Goya a la mejor película de animación y una cinta dirigida a cuatro manos entre Fernando Trueba y Javier Mariscal. Narra una historia de amor entre un pianista y una cantante en la Cuba de finales de los años 40.



↓ Ver sesiones

• Género: Animación
 • Director: Fernando Trueba, Javier Mariscal
 • Intérpretes: Eman Xor Oña, Estrella Morente, Limara Meneses y Mario Guerra
 • Nacionalidad y año de producción: Cuba, España
 • Duración: 84 min
 • Fecha de estreno: 25 / 02 / 2011
 • Productora: Isle of Man Film, CinemaNX
 • Guionista: Fernando Trueba
 • Distribuidora: Walt Disney Studios Motion Pictures Spain
 • Música: Bebo Valdés
 • Calificación: todos los públicos

• Guíadelocio.com: ★★☆☆☆
 • Valoración usuarios: ★★★★★ (29 votos)
 • Vota: ★★★★★

Sinopsis Sesiones Artículos Criticas Trailers Galería Comentarios

Sinopsis

Los años 40. La Habana. Chico es un pianista que sueña con convertirse, algún día, en una figura importante del jazz. Rita es una hermosa cantante que, con su exquisita y dulce voz, se lleva los mayores aplausos por parte del público que abarota la sala en la que actúa. Un día, sus destinos se cruzan y se enamoran. Juntos comienzan una relación de desgaste, llena de altibajos emocionales y pasiones, que les permitirá vivir su profesión desde Nueva York a Hollywood al son de un bolero.


Chico y Rita es una película de animación que supone la colaboración del dibujante Javier Mariscal y el director Fernando Trueba (*Belle Époque*, *El baile de la victoria*), en un intento personal por hacer un particular homenaje al mundo del cine y de la música. Se trata de una historia de amor clásica de encuentros y desencuentros, muy bien amparada por el son cubano, una música de la que Trueba siempre ha sido un ferviente admirador, como ya demostró en la lograda *Calle 54*.

La cinta muestra la época dorada del jazz, esos años 40 en los que se comenzaba a experimentar con otros sonidos, dando origen a nuevos ritmos como el jazz latino. Los actores de teatro Eman Xor Oña y Limara Meneses prestan sus voces al pianista Chico y a la sensual Rita, respectivamente. En el apartado vocal también interviene la cantante granadina Estrella Morente. Chico y Rita, que tiene una estética de cómic gracias a los fotogramas diseñados por el propio Mariscal, ganó el Goya a la mejor película de animación.

Comentarios

- c.etcheverry@coac.es (1 de 1).
Supongo que uno puede ir al cine por distintos motivos, y también con distintas intenciones. Es posible que un crítico lo haga para estudiar y ampl...

Galería



A continuación mostramos un extracto de la estructura de la página web de una película, explicando sus diferentes partes y cómo es tratado por el parseador de HTML Jericho.

Extractos de código de la página de una película:

```
<h1class="nobgftl">
  Chico y Rita
</h1>
```

Buscando en el código de la página de la película el elemento con el identificador “nobgftl” obtenemos el título de la película

```
<div class="infoContent">
  <ul>
    <li>Género: <!-- ... -->Animación.<!-- ... --></li>
    <li>Director: Fernando Trueba, Javier Mariscal</li>
    <li>Intérpretes: EmanXor Oña, Estrella Morente, Limara Meneses y Mario
    Guerra.</li>
    <li>Nacionalidad y año de producción: Cuba, España</li>
    <li>Duración: 94 min.</li>
    <li>Fecha de estreno: 25 / 02 / 2011.</li>
    <li>Productora: Isle of Man Film, CinemaNX.</li>
    <li>Guionista: Fernando Trueba.</li>
    <li>Distribuidora: Walt Disney Studios Motion Pictures Spain.</li>
    <li>Música: Bebo Valdés.</li>
    <li class="last">Calificación: Todos los públicos<!-- ... --></li>
  </ul>
</div>
```

Para obtener la información técnica sobre la película procesamos el fragmento de código de la clase “infoContent”. Una vez lo tenemos, vamos buscando si, en los elementos de la lista que contiene, aparecen los identificadores de los campos de la clase Cine (“Director”, “Duración”, “Intérpretes”,...), almacenando aquellos que aparezcan.

```
<div class="wysiwyg">
  <p class="sep01">Años 40, La Habana. Chico es un pianista que
  sueña con convertirse, algún día, en una figura importante del jazz. (...)
  <strong>Chico y Rita</strong>, que tiene una estética de cómic gracias a
  los fotogramas diseñados por el propio Mariscal, ganó el Goya a la mejor
  película de animación.</p>
</div>
```

Para almacenar la sinopsis de la película solo tenemos que buscar el elemento de la clase “wysiwyg”.

```
<div class="filshadow">
  
</div>
```

Por último, para descargarnos el cartel de la película solo tenemos que acceder al primer elemento de tipo img dentro del código de la clase “filshadow”.

4.1.6.2 EXTRACCIÓN E INFORMACIÓN DE CINES Y PASES

Encargado de extraer toda la información de cada cine y los pases de cada película.

4.1.6.2.1 Información extraída por cada cine

- URL del cine dentro de La Guía del Ocio: Elemento que utilizaremos (al igual que ya hicimos con la URL de cada película) como identificador de cada cine antes de introducirlo en la base de datos.
- Nombre
- Provincia
- Número de salas
- Teléfono
- Web oficial
- Ciudad
- Dirección
- Zona
- Parking
- Dirección web de compra de entradas para ese cine
- Lista de los pases de películas de ese cine.

4.1.6.2.2 Información extraída por cada pase

- URL de la película del pase dentro de La Guía del Ocio
- Día
- Hora

4.1.6.2.3 Funcionamiento detallado

Al igual que hicimos con el extractor de películas, cargaremos inicialmente una página que contenga todos los cines de una provincia (<http://www.guiadelocio.com/salas/<PROVINCIA>>), de la que extraeremos los enlaces a la ficha de cada cine. A continuación deberemos explorar las páginas web de cada cine para extraer la información deseada.

Página de los cines de una provincia:

guía del ocio .com

en Guíadelocio.com en Madrid en cine

¿Qué buscas en Guíadelocio.com? » ¡Encuentra!

Ocio en tu zona Club GDO

Madrid Ver otras provincias Seleccionar localidad

Portada Cine Restaurantes Teatro y danza Conciertos Arte Tarde y noche Niños Ferias Despedidas y fiestas Cursos y aficiones

Estás en: Ocio en tu zona > Madrid > Cine

Síguenos en: f t

Cine en Madrid

Estrenos
Cartelera
Críticas
Trailers
Próximos estrenos
Estrellas
Ranking
Galerías
Personajes
Salas
Noticias
Archivo películas

Buscar por:
• Género

- Todos
- Acción (94)
- Animación (56)
- Aventuras (20)
- Comedia (204)
- Documental (84)
- Drama (297)
- Fantástica (7)
- Romántica (28)
- Terror (32)
- Thriller (59)

Película: Selecciona Cine: Selecciona Ir →

Salas

Madrid

- Cine Estudio (Círculo de Bellas Artes)**
Calle Alcalá, 42
- Abaco Villaverde**
Avenida Andalucía, km 7.
Agua para elefantes, Arthur y la guerra de los mundos, Brutal Box, El inocente (2011), El sicario de Dios, El último exorcismo, Fast and Furious 5, Hop, Piratas del Caribe: En mareas misteriosas, Río, Thor, X-Men: Primera generación,
- Acteón**
Calle Montera, 31
Agua para elefantes, Brutal Box, El árbol, El sicario de Dios, Midnight in Paris, Pequeñas mentiras sin importancia, Piratas del Caribe: En mareas misteriosas, ¡Qué dilema!, Sin identidad, X-Men: Primera generación,
- Artesiete Alcalá Full HD**
Calle Alcalá, 414
Fast and Furious 5, Midnight in Paris, Pequeñas mentiras sin importancia,

Club GDO

Únete al Club Guía del Ocio
Gana muchos premios
¡Entra! ¡Hazte miembro!

FUELEIDAD: FUELEIDAD:

Lo +

Lo + valorado Lo + comentado Lo + reciente

- 1 Retornos
Valoración: ★★★★★ (5 votos)
- 2 Carl Gustav Jung
Valoración: ★★★★★ (4 votos)
- 3 Alexia
Valoración: ★★★★★ (4 votos)
- 4 Recién graduada
Valoración: ★★★★★ (3 votos)
- 5 Amanecer de un sueño
Valoración: ★★★★★ (3 votos)

Personajes

Enlaces a las páginas con las fichas de los cines


```

<div class="moduleType101 borderb clear">
  <pclass="texto withBolo01">
    <a href="/cine/madrid/madrid/acteón">
      <strong>Acteón</strong><br/>
    </a>Calle Montera, 31 <br/>
    <a href="/cine/archivo-peliculas/agua-para-elefantes">
      <strong>Agua para elefantes</strong>
    </a>,
    <a href="/cine/archivo-peliculas/brutal-box">
      <strong>Brutal Box</strong>
    </a>,
    ...
    <a href="/cine/archivo-peliculas/x-men-primera-generacion">
      <strong>X-Men: Primera generación</strong>
    </a>,
  </p>
</div>

```

Cada división de la clase “moduleType101 borderbclear” en el código de la página contiene un cine (el enlace a su ficha y los enlaces a las películas que proyecta).

Página con la ficha de un cine:

Zona

Dirección

Información

Enlace de referencia

Enlace de la película


```

<div class="titlcon type11">
  <div>
    <h1>Cine en Madrid, Madrid Capital</h1>
  </div>
</div>

```

Para obtener la zona del cine, solo tenemos que buscar la división de la clase "titlcon Type11"

```

<div class="moduleType106 clear">
  <div class="content clear">
    <ul>
      <li>Dirección: Calle Montera, 31 , 28013, Madrid</li>
    </ul>
    <div id="map_canvas" style="width:305px; height:289px;">
    </div>
  </div>
</div>

```

La dirección del cine se encuentra en el primer elemento de la lista en la división de clase "moduleType106 clear"

```

<div class="infoContent" style="margin-left:0">
  <ul>
    <li>Sábados, domingos y festivos (Vísperas y festivos): 8 euros €</li>
    <li>Laborables (Laborables): 7,5 euros €</li>
    <li>Sábados, Domingo, Festivos (Sesion antes de las 17h): 6 euros €</li>
    <li>Sesiones antes de las 17,00h (Primera sesión): 5,5 euros €</li>
    <li>Acceso: Gran Vía Metro Gran Vía ; Plaza del Sol Metro Sol</li>
    <li>Día del espectador: Miércoles 5,5 €</li>
    <li>Número de salas:9</li>
    <li>Parking:No</li>
    <li>Teléfono:34-915.22.22.81</li>
    <li>Web:<a href="http://www.filmofono.com">www.filmofono.com</a></li>
    <li>Venta de entradas:<a href="http://www.entradas.com/entradas/a001006.do?identidad=1&id canal=2&idcine=22&idprov=28">http://www.entradas.com/entradas/a001006.do?identidad=1&idcana l=2&idcine=22&idprov=28</a>
    </li>
  </ul>
</div>

```

De forma similar a la utilizada para la obtención de la información técnica de las películas obtenemos los campos de información sobre el cine, contenidos en una lista a la que podemos acceder buscando en el código la división de la clase "infoContent"

```

<div class="contentSesiones">
  <ul class="dias" style="width: 80%;">
    <li class="desactive" style="padding: 5px 10px;">

```

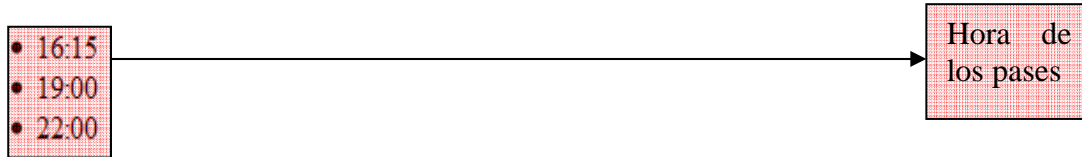
```

        <div class="izq"></div>
        viernes 3
        <div class="der"></div>
    </li>
    ...
    <li class="active" style="padding: 5px 10px;">
        <div class="izq"></div>
        <a
            href="/ezjscore/call/sesiones::pelicula::31321::645316::1307397600">m
            artes 7</a>
        <div class="der"></div>
    </li>
    ...
    <li style="padding: 5px 10px;">
        <div class="izq"></div>
        <a
            href="/ezjscore/call/sesiones::pelicula::31321::645316::1307570400">ju
            eves 9</a>
        <div class="der"></div>
    </li>
</ul>
<div class="horaSesiones sep01 clear">
    <div class="dirSesion">
        <p class="withBolo01">
            <a href="/cine/archivo-peliculas/agua-para-elefantes">
                <strong>Agua para elefantes</strong>
            </a>
        </p>
    </div>
    <div class="diasSesion">
        <div class="horas clear">
            <ul class="ftl" style="width:100%">
                <li class="ftl">16:15</li>
                <li class="ftl">19:00</li>
                <li class="ftl">22:00</li>
            </ul>
            <span class="boton02 sep05">
                <a
                    href="http://www.entradas.com/entradas/a001006.do?identidad=
                    1&idcanal=2&idcine=22&idprov=28&affiliate=105"
                    target="_blank">Comprarentradas</a>
            </span>
        </div>
    </div>
</div>
</div>
</div>

```

Para obtener los pases de las películas el proceso se complica ligeramente. Cada película proyectada en el cine tiene todos sus pases contenidos en un divisor de la clase “contentSesiones”. Ha de comentarse que las horas de los pases de un día se encuentran en un archivo independiente de este, accesible por el enlace que hay en cada película y por cada día. Puesto que puede que algunos enlaces no estén disponibles (cada día desactivan los enlaces del día anterior), lo que hacemos es buscar el enlace del último día de la lista (jueves), que siempre estará activo, guardarlo, y calcular a partir de él los de los días anteriores. Una vez los tenemos, vamos abriendo los enlaces para guardar las horas de los pases de ese día.

Página con las horas de los pases de un día:

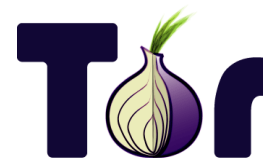


```

<div class="horas clear" style="padding-top: 0px;">
  <ul class="ftl" style="width:100%">
    <li class="ftl">16:15</li>
    <li class="ftl">19:00</li>
    <li class="ftl">22:00</li>
  </ul>
</div>
  
```

De la página de las horas de los pases nos quedamos con todos los elementos de la lista

4.1.7 TOR



Debido a la extracción de información de la web de “La Guía del Ocio”, se accede a un volumen enorme de páginas web. Esta acción se lleva a cabo semanalmente, lo que implica un tráfico de información elevado. Si tenemos en cuenta el número de pruebas realizadas con el extractor de información, alcanzamos cuotas de acceso que podrían alarmar al administrador de “La Guía del Ocio”. Éste podría sospechar que se está realizando un uso indebido de su página web e incluso tomar acciones de restricción de acceso a su web desde la dirección IP asociada al extractor, impidiéndonos continuar con nuestra labor.

Como consecuencia de estas preocupaciones, y a modo preventivo, decidimos incorporar un sistema de comunicación anónima para evitar ser localizados y que se nos prohibiese acceder a la información. TOR nos permite alcanzar nuestro objetivo de un modo relativamente sencillo.

TOR (TheOnionRouter) provee un canal de comunicación anónimo y permite librarse de la detección por análisis de tráfico. TOR, en un principio, garantiza el anonimato, pero la información que circula a través de los encaminadores puede ser vista por terceras personas.

TOR está incluido en el extractor de información gracias a la ayuda de la librería desarrollada por silvertunnel.org llamada Netlib, y puede ser activado y desactivado a voluntad del administrador de la aplicación. En un principio no hemos tenido que recurrir a su uso, ya que no se nos ha baneado ni advertido.

El único problema de su uso es la eficiencia temporal, que se ve ampliamente perjudicada. El extractor de información dura un tiempo mucho mayor para realizar su tarea cuando se está usando navegación anónima. Aun así, al ser una tarea de periodicidad semanal, el tiempo no supone un precio demasiado alto a pagar.

4.2 Categorizador de géneros

Como se puede ver en el apartado de conocimiento, la recomendación se va a basar principalmente en la distribución de géneros por temperamentos. Para ello disponemos de los datos que se han podido extraer de página web de “La Guía del Ocio”, que deben ser tratados de algún modo para que se adapten a nuestro modelo.

Por un lado tenemos el género de la película tal y como viene especificado en “La Guía del Ocio”. Esta información de adapta correctamente a nuestra idea principal de recomendación, pero no es suficiente. Si tenemos en cuenta únicamente el género que viene por defecto, no podríamos dar más que once puntuaciones distintas por usuario, es decir, una puntuación por género, al suponer que las películas son 100% del género que viene especificado en la web). Obviamente, si basamos el resultado únicamente en este dato, las puntuaciones no van a tener precisión suficiente, y se van a encontrar recomendaciones con muchos empates.

Analizando cualquier película, podemos ver que realmente no pertenece a un único género, sino que se trata de una composición de varios géneros distribuidos en mayor o menor medida (por ejemplo, una película puede ser un 60% Acción, 30% Humor y 10% Romántica). La manera ideal de categorizar por géneros sería viendo una a una todas las películas que hay en cartelera y analizándolas detalladamente, de modo que pudiésemos extraer sus porcentajes de género. Esta tarea es inviable, así que se debe buscar un modo de automatización de este proceso (usando razonamiento basado en casos, CBR).

El CBR a día de hoy es ampliamente usado, así como uno de los métodos más exitosos dentro del campo de la inteligencia artificial. Para la implementación de un CBR, primeramente hace falta un conjunto de casos resueltos que sirvan de base de conocimiento. A partir de esa base de casos, y aplicando funciones de similitud, podemos llegar a la solución de nuevos casos (es un modelo basado en el razonamiento natural que realiza el ser humano). Para nuestro proyecto vamos a centrarnos en el CBR textual con lenguaje natural, que consiste en un subapartado del CBR en el que la base de conocimiento está en formato texto.

Como se ha comentado anteriormente, aparte del género propio con el que categorizan las películas en la web de “La Guía del Ocio” contamos también con la sinopsis. La sinopsis resume el argumento de la película en cuestión, mostrando su esencia. Un ser humano podría leer una sinopsis y a partir de ella intentar decidir cuáles son los géneros de esa película. Esta idea es la que se desea extrapolar a nuestro trabajo mediante el uso de CBR textual.

El proceso planteado se va a llevar a cabo mediante el uso de JCOLIBRI2. A grandes rasgos podríamos decir que JCOLIBRI es un proyecto llevado a cabo por el grupo de investigación GAIA de la UCM. Se trata de un Framework que facilita la implementación de aplicaciones CBR. En concreto JCOLIBRI2 está orientada a los desarrolladores (mientras que la primera versión consta de una interfaz gráfica y está orientada para usuarios).

Siguiendo los pasos explicados, primeramente tenemos que crear una base de casos resueltos formada por un conjunto de películas que contengan sinopsis y género, y que sirvan para la resolución de los nuevos casos planteados (las películas que vayan a encontrarse en un futuro en la cartelera). La decisión tomada ha sido la de rellenar la base de casos con películas que sean altamente representativas de un género, tanto por su calidad como por su contenido. Por ejemplo, “El Exorcista” como uno de los referentes del cine de terror, “Fargo” como un referente de Thriller; y otras con dos géneros como “Amelie” como representación de comedia romántica. La lista completa de películas base se puede encontrar en el proyecto como “películas.txt”. Es una lista sencillamente modificable e incluso intercambiable con otras listas siempre que sigan el siguiente formato:

```

<pelicula>
    Título de la película 1
<genero>
    Género 1
    Género 2
    ...
    Género n
<sinopsis>
    Sinopsis de la película
<pelicula>
    Título de la película 2
<genero>
    Género 1
    Género 2
    ...
    Género n
<sinopsis>
    Sinopsis de la película 2
    ...
<pelicula>
    Título de la película n
<genero>
    Género 1
    Género 2
    ...
    Género n
<sinopsis>
    Sinopsis de la película n

```

Vemos que la lista puede tener un número indefinido de películas, que deben empezar por la etiqueta <pelicula>, tras la que se debe poner el título. Luego se debe colocar la etiqueta <genero>, y debajo de ella los géneros pertenecientes a la película en cuestión, siempre y cuando se ajusten a los once géneros clave que usamos en nuestra base de conocimiento. Por último está la etiqueta <sinopsis>, tras la que se debe encontrar la sinopsis de la película en una sola línea.

Una vez dada por completa la base de conocimiento, mediante los algoritmos facilitados por JCOLIBRI, podemos pasar a categorizar nuevas películas extrayendo de ellas sus porcentajes de género. Ya que el texto de la sinopsis viene en lenguaje natural en español, debe someterse previamente a un tratamiento de limpieza de palabras vacías, extracción de raíces, etc. Para este proceso nos servimos de paquetes que funcionan correctamente en JColibri tales como OpenNLPSpanish (procesador de lenguaje natural), StopWordsDetectorSpanish (que elimina signos de puntuación o palabras vacías) y TextStemmerSpanish (que determina las raíces de las palabras).

Una vez ha sido tratado el texto de la sinopsis, podemos pasar a la fase de recuperación de casos con el algoritmo LuceneSpanish. Este algoritmo es usado para la búsqueda de casos a partir de la similitud de atributos textuales y está basado en el Modelo del Espacio Vectorial, donde la similitud entre casos es medida en base al ángulo formado por sus representaciones vectoriales. Es un sistema que ofrece buenos resultados y que se ejecuta en muy pocos segundos para todas las películas que hay en cartelera.

El proceso de categorización de una película en porcentajes de géneros se lleva a cabo a partir de los resultados obtenidos por el algoritmo KNN, que devuelve los “k” casos recuperados más similares al que deseamos categorizar. Los vecinos recuperados vienen acompañados de un valor comprendido entre 0 y 1, que indica el grado de similitud de la recuperación respecto al caso de consulta.

En las ecuaciones siguientes se explica el procedimiento seguido para el cálculo de porcentajes de géneros de una película. En ellas tenemos casos recuperados que tienen un valor de similitud (siendo cada caso recuperado un conjunto de géneros) y géneros que aparecen en los casos recuperados.

$$\forall \text{género } g, \forall \text{ caso } C \text{ t. } q. C \in \{\text{casosRecuperados}\} \wedge g \in C$$

$$\text{valor acumulado } g \equiv w_g = \sum_{C=\text{caso recuperado } 1}^{\text{caso recuperado } k} \frac{\text{similitud}(C)}{|C|}$$

$$\text{porcentaje } g = \frac{w_g}{\sum_{i=1}^{\text{géneros recuperados}} w_i}$$

*Siendo g un género de película y
C un caso recuperado.*

El proceso de categorización se debe realizar una vez a la semana, ya que cada semana hay películas nuevas de las que se debe extraer su conjunto de géneros. Es un proceso que dura pocos segundos, así que no hay problema en renovar todas las categorías semanalmente. Las películas y su información se extraen de las tablas correspondientes de la base de datos, se analizan, y se suben sus porcentajes de géneros a la tabla que contiene películas y géneros. Cabe destacar que las películas que no dispongan de sinopsis no serán categorizadas, su distribución será del 0% para todos los géneros, implicando que su recomendación posterior valga 0.

4.2.1 Ejemplo de resultados

4.2.1.1.1 Título: X-men: Primera Generación

SINOPSIS:

El joven Charles Xavier posee una habilidad que le distingue de los humanos, puede leer el pensamiento e influir sobre otras personas. Se une a otro chico excepcional llamado Erik Lenhsherr, con el poder de alterar el campo magnético y cambiar a voluntad los metales. Ambos son mutantes y juntos, tratan de llevar a su especie evolucionada a coexistir con los humanos. Las discrepancias que se crearan entre ellos, dará lugar a una feroz lucha entre el creador de la patrulla X y Magneto (continúa...).

DISTRIBUCIÓN GÉNEROS:

- Acción: 17'2865%
- Aventura: 24'6891%
- Drama: 8'9169%
- Thriller: 8'0244%
- Ciencia Ficción: 41'0831%

4.2.2 Problemas y errores. Soluciones

Uno de los principales problemas encontrados es que la Animación y el Documental son géneros que no se pueden obtener a partir de la sinopsis. Leyendo el argumento de una película no tiene por qué distinguirse la pertenencia a uno de estos dos géneros. La solución encontrada ha sido la de crear un caso recuperado ficticio, es decir, a parte de los casos recuperados por KNN, se añade un caso más con similitud 1 y con el género propio de la película. De este modo, añadimos con bastante peso el género establecido por “La Guía del Ocio”. El peso del caso sintético se verá modificado a partir del resto de casos recuperados, ya que si estos tienen grado de similitud muy alto el resultado será una distribución bastante equitativa, mientras que si el grado de similitud es bajo el resultado será una distribución de géneros donde tendrá un peso mucho mayor el género definido del caso sintético.

Esta solución es muy buena no solo para solventar el problema de los géneros Animación y Documental, sino que sirve para cualquier tipo de género ayudando a distribuir los porcentajes de un modo en el que se elimina mucho ruido que pudiesen traer recuperaciones de baja calidad.

Un error posible es que los géneros que encontremos en la recuperación de información de la web no se adapten a los géneros que usamos habitualmente en los algoritmos del recomendador (por ejemplo, la aparición de los géneros Docudrama, Western, Familiar, etc.). La mayoría de casos han sido contemplados, pero en caso de que en la web añadieran un nuevo género se debería solventar el problema en la implementación de la clase Géneros. En esa clase se añadiría manualmente el género nuevo, asociándolo a uno de los once géneros existentes a partir del criterio del desarrollador. En caso de no querer añadirlo, el categorizador seguirá funcionando, pero en esa película no añadirá un caso de similitud máxima con el género propio del filme.

El último error encontrado es el de que un editor de la web haya escrito mal la sinopsis y cometido un error de puntuación, paréntesis desparejados, etc. En este caso se debe modificar la sinopsis de la tabla que contiene la sinopsis de la base de datos de modo manual, ya que de otro modo el algoritmo de tratamiento de lenguaje natural es incapaz de extraer el contenido correctamente.

Estos dos últimos errores son poco comunes, pero pueden darse en algún momento. Por ese motivo creemos conveniente comentarlos y aportar la solución.

5 SISTEMA RECOMENDADOR

Si el conocimiento es la base sobre la que se asienta el proyecto, el sistema recomendador es el motor encargado de hacerlo funcionar, la parte encargada de gestionar la información de la que disponemos, trabajar con ella y obtener los resultados deseados.

Se debe encontrar la manera en la que conseguir recomendar una película a un usuario concreto, a partir de las relaciones que hemos establecido entre los cuatro temperamentos de Keirsey y los géneros cinematográficos contemplados. Con el recomendador se lleva a cabo esta tarea, desde diversos enfoques, ayudando al usuario a elegir qué película debería ir a ver en los cines de su ciudad.

El esqueleto del sistema recomendador está formado por distintas piezas, distintos recomendadores enfocados desde planteamientos e ideas completamente diferentes. Por tanto, podríamos decir que contamos con tres recomendadores implementados por separado, los cuales, se unifican en un único recomendador de modo que el resultado del trabajo elaborado por cada uno de ellos individualmente se combina para llegar a una solución mejor.

De manera visual, el siguiente gráfico ilustra a grandes rasgos el diseño de la aplicación.



5.1 Recomendador Temperamento/Género o RecomendadorMix

Este recomendador es el primero de los tres implementados. Se basa en la idea más intuitiva de todas; la de usar directamente la tabla de conocimiento combinada (tabla 3), ya que tanto el enfoque del recomendador como la tabla trabajan con los dos datos a la vez (temperamentos y géneros).

El usuario está representado por una distribución de temperamentos conseguida a partir del test de Keirsey. Los temperamentos, como ya se ha explicado en la parte de conocimiento, los podemos encontrar en las filas de la tabla3, sabiendo que el contenido de la tabla es una repartición de pesos entre los once géneros tenidos en cuenta y que se encuentran situados en las columnas de la tabla.

La idea consiste en calcular a partir de una película concreta (distribución de géneros en porcentajes, por ejemplo, 70% Acción y 30% Aventura) y un usuario concreto (distribución de temperamentos en porcentajes, por ejemplo, 30% Artesano, 40% Guardián, 10% Idealista y 20% Racional) la puntuación obtenida al combinarlos. Es decir, con la llegada de una tupla formada por Usuario y Película se calcula su evaluación.

Para ello, se multiplica el porcentaje de cada temperamento del usuario, por el contenido que se encuentra en la tabla3 por cada género de los que componen la película y por el porcentaje que ocupa ese género en la película. Con la siguiente fórmula se puede ver en detalle.

$$\begin{aligned} & \text{puntuación}(\text{Usuario}, \text{Película}) \\ & \quad \text{Racional} \quad \text{Ciencia Ficción} \\ & = \sum_{t=\text{Artesano}} \sum_{g=\text{Acción}} \text{Usuario}(t) * \text{Película}(g) * \text{tabla3}(t, g) \end{aligned}$$

Siendo $\text{Usuario}(x)$ un porcentaje de temperamento,

$\text{Película}(y)$ un porcentaje de género y

$\text{tabla3}(x, y)$ el contenido de la tabla en esa posición.

Con esto conseguimos la puntuación de una película para un usuario. Ahora simplemente se debe repetir este cálculo para todas las películas que queramos. De los resultados obtenidos, si los ordenamos de mayor a menor puntuación, podemos extraer la lista de películas recomendadas para este usuario concreto.

Diremos a modo introductorio que los resultados para este recomendador son muy buenos. Para más detalle sobre el análisis de resultados se puede consultar la sección de “Análisis de pruebas”.

5.2 Recomendador Género/Género Global

Para este segundo recomendador cambiamos la perspectiva usada en el anterior. Volvemos a recibir un usuario que sigue siendo una distribución de temperamentos, pero en esta ocasión la intención es transformar a este usuario en una distribución de géneros (su reparto ideal de géneros) y calcular la similitud existente entre su asignación de géneros y la distribución de los géneros que forman la película.

Primeramente debemos obtener la distribución de géneros del usuario. Para ello, por cada género de los existentes multiplicaremos los porcentajes de temperamento del usuario por el contenido de la tabla3 para la posición de ese temperamento con el género que se esté tratando. Se puede ver en la siguiente fórmula:

$$\forall \text{Género } g, \\ \text{Usuario}(g) = \text{cantidad}(\text{Usuario}, g) = \sum_{t=\text{Artesano}}^{\text{Racional}} \text{Usuario}(t) * \text{tabla3}(t, g)$$

Una vez obtenida la repartición de géneros ideales para un usuario concreto, ya podemos realizar la comparación del usuario respecto a una película. La idea es buscar la similitud que hay entre Usuario y Película basándonos en los géneros. Se calcula la media de separaciones que hay por cada género, y se resta a la unidad. De este modo, cuando una película tenga mucha similitud a un usuario su puntuación se acercará al valor “1” mientras que, en caso contrario, una gran diferencia entre película y usuario recortará la puntuación a valores que giran por encima de valores entorno al “0”.

Podemos ver la idea en la siguiente fórmula:

$$\text{similitud}(\text{Usuario}, \text{Película}, \text{Género}) = 1 - \frac{|\text{Usuario}(\text{Género}) - \text{Película}(\text{Género})|}{100} \\ \text{puntuación}(\text{Usuario}, \text{Película}) = \frac{\sum_{g=\text{Acción}}^{\text{CienciaFicción}} \text{similitud}(\text{Usuario}, \text{Película}, g)}{|\text{Géneros}|}$$

Siendo Géneros el conjunto que engloba todos los géneros.

Una vez conocida la función de puntuación, lo único que debemos hacer es iterarla con todas las películas, de modo que luego podamos ordenarlas en orden decreciente en base a su puntuación para obtener la lista de recomendación.

Este recomendador por lo general da peores resultados que el anterior, pero existen casos en los que funciona muy bien. Para más detalle consultar la sección de “Análisis de pruebas”.

5.2.1.1 RECOMENDADOR GÉNERO/GÉNERO LOCAL

Este recomendador es una modificación del anterior, es decir, del Género/Género global. La idea principal es la misma, pero con ligeros cambios en algunos conceptos.

En el global se trataban todos los géneros cinematográficos, tanto para establecer la distribución de géneros ideal para un usuario como para calcular la puntuación.

Si tenemos en cuenta, que una película no tiene todos los géneros existentes sino que consiste en una distribución de un subconjunto de géneros, tiene sentido plantear la posibilidad realizar los mismos cálculos que antes pero teniendo en cuenta únicamente los géneros concretos de los que dispone la película a puntuar.

De este modo, la distribución de géneros para el usuario se realizaría de modo aislado para cada película usando esta fórmula. Es similar el proceso anterior, pero cambia la precondition:

$$\forall \text{Género } g \quad t.q. \quad Película(g) > 0$$

$$Usuario(g) = cantidad(Usuario, g) = \sum_{t=Artesano}^{Racional} Usuario(t) * tabla3(t, g)$$

Una vez obtenida la distribución ideal de géneros del usuario para los géneros que forman parte de la película en cuestión, se calcula la puntuación. Vuelve a ser un proceso similar al anterior, pero limitando los géneros que se tratan. Lo vemos en las siguientes fórmulas:

$$similitud(Usuario, Película, Género) = 1 - \frac{|Usuario(Género) - Película(Género)|}{100}$$

$$puntuación(Usuario, Película) = \frac{\sum_{g=Acción}^{CienciaFicción} similitud(Usuario, Película, g)}{géneros(Película)}$$

Donde la función $géneros(Película)$ devuelve en este caso el número de géneros que forman parte de la película.

Si iteramos los cálculos de puntuación para todas las películas, otra vez podemos obtener la lista de recomendaciones para el usuario.

El problema reside, en que en esta ocasión, los resultados distan ampliamente de lo esperado. Al realizar algunas pruebas se comprueba fácilmente que las recomendaciones realizadas se apartan de lo que dicta la lógica impuesta de las tablas de conocimiento.

Este error se debe a que, para todo subconjunto de géneros, existe una distribución ideal de esos géneros para el usuario. Si la distribución de esos géneros en la película es muy parecida a la calculada para el usuario, esa película obtendría una puntuación elevada, pero sin tener en cuenta si realmente esos géneros le gustan al usuario o no. Es decir, una película con dos géneros que un usuario pueda odiar, si se distribuyen de cierto modo, damos con una combinación ideal de esos dos géneros para el usuario en cuestión, pero sin tener en cuenta que, realmente, odiaría esa película. Si pusiésemos a su lado una película de dos géneros que encantan al usuario, pero que por desgracia se han distribuido de un modo ligeramente distinto a la combinación ideal de esos dos géneros para ese usuario, resultaría una puntuación menor que la obtenida por el recomendador para la película odiada.

Por tanto, tras analizar los resultados obtenidos, se descarta el uso de este recomendador en el proyecto.

5.3 Recomendador Temperamento/Temperamento

El último recomendador implementado para el proyecto es el recomendador de Temperamento/Temperamento. El concepto es similar al planteado en el recomendador basado en géneros, pero contemplado desde el punto de vista opuesto. En esta ocasión, en lugar de convertir una distribución de temperamentos de usuario en una lista de géneros, lo que haremos será seleccionar los géneros correspondientes a una película y transformarlos en una distribución de temperamentos.

Para convertir un conjunto de géneros en un reparto de temperamentos, volveremos a usar la tabla3 como fuente de conocimiento. Recibiremos una película con su correspondiente categorización de géneros, y mediante los cálculos que se muestran a continuación extraeremos el temperamento del filme:

$$\forall \text{Temperamento } t, \\ \text{Película}(t) = \text{cantidad}(\text{Película}, t) = \sum_{g=\text{Acción}}^{\text{CienciaFicción}} \text{Película}(g) * \text{tabla3}(t, g)$$

El objetivo es, una vez obtenido el temperamento de la película, poder aplicar una función de similitud que nos permita comparar la película con el usuario. Cuanto más parecido guarden los dos elementos (calculando la media de los parecidos), mayor será la puntuación obtenida. Nuevamente, son fórmulas parecidas a las usadas en el recomendador de géneros, pero con pequeños cambios en las variables, tal y como podemos observar:

$$\begin{aligned} &\text{similitud}(\text{Usuario}, \text{Película}, \text{Temperamento}) \\ &= 1 - \frac{|\text{Usuario}(\text{Temperamento}) - \text{Película}(\text{Temperamento})|}{100} \end{aligned}$$

$$\text{puntuación}(\text{Usuario}, \text{Película}) = \frac{\sum_{t=\text{Artesano}}^{\text{Racional}} \text{similitud}(\text{Usuario}, \text{Película}, t)}{|\text{Temperamentos}|}$$

Siendo Temperamentos el conjunto de cuatro temperamentos de Keirsey.

De este modo tenemos la puntuación calculada para la pareja formada por película y usuario. Nuevamente, debemos recorrer la lista de películas, distribuyendo en cada una su proporción de temperamentos a partir de su colección de géneros y calculando la puntuación. Ordenadas decrecientemente por puntuaciones tenemos la lista de recomendación otorgada por este recomendador.

Este recomendador ofrece buenos resultados, así que también lo tendremos en cuenta a la hora de juntar los recomendadores para llegar a la recomendación final. Para más detalle sobre los resultados, mirar la sección de “Experimentos y Resultados”.

6 APLICACIÓN WEB PARA FACEBOOK

6.1 Introducción

Para hacer nuestra aplicación accesible optamos por desarrollar una aplicación en Facebook. Facebook permite que aplicaciones web puedan acceder a los datos de sus usuarios (siempre que estos lo permitan), evitándose así el desarrollo de un sistema de registro. Además facilita el acceso a la aplicación y su difusión, siendo esto factores importantes para potenciar el uso de una aplicación.

6.2 Registro en Facebook

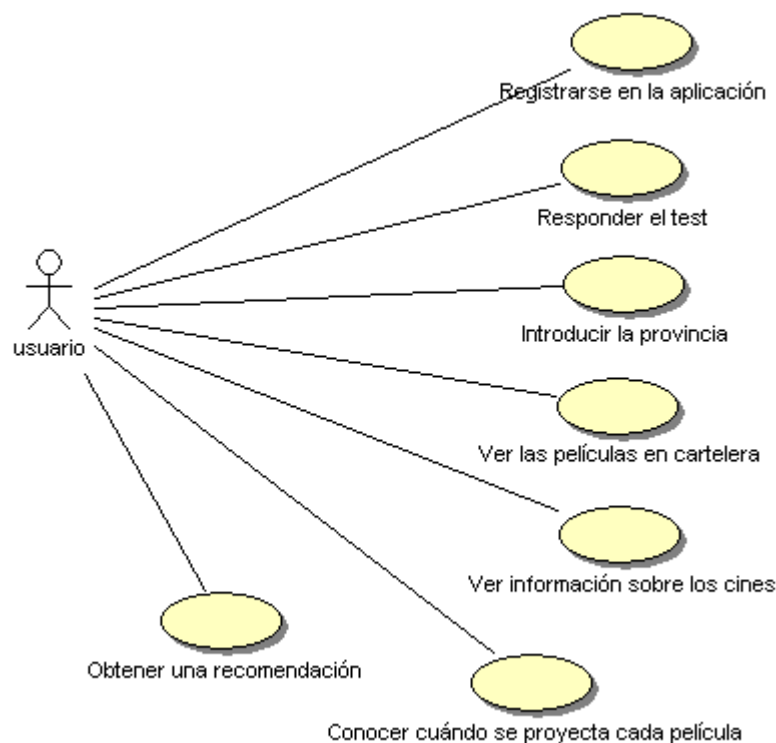
Para poder desarrollar una aplicación en Facebook es necesario tener una cuenta de usuario y estar dado de alta en la aplicación Developers (www.Facebook.com/developers). En ella encontramos documentación sobre la creación de aplicaciones, al igual que el acceso al registro de aplicaciones nuevas. Una condición necesaria para el registro de una aplicación es que la cuenta de Facebook desde la que se quiere hacer esté verificada (método que se puede realizar introduciendo un número de teléfono que pueda recibir un código de confirmación por sms o una cuenta bancaria). Una vez tenemos todo esto, podemos crear una aplicación pulsando el botón “crear una nueva aplicación”. Tras seguir los pasos que nos propone Facebook, llegaremos al panel de control de la aplicación:

The screenshot shows the Facebook Developers application dashboard for an application named 'Moodvie'. At the top, there is a navigation bar with 'Mis aplicaciones' and a button '+ Crea una nueva aplicación'. The main content area has a sidebar on the left with the application icon and name 'Moodvie'. The main panel displays the application's status as 'Estado del directorio: No enviado' and provides statistics: 'Usuarios activos por mes' (85), 'Personas a las que les gusta' (1), and 'Total de usuarios' (60). Below these statistics, there is a list of application details including the application ID (148131101921752), API key (ea878b99a4ae3829d65855adfa64c), secret code (0 [redacted] b), site URL (http://supergaia.fdi.ucm.es:8060/moodvie/), page URL (http://apps.facebook.com/moodvie/), canvas URL (http://supergaia.fdi.ucm.es:8060/moodvie/), contact email (moodvie@hotmail.es), support email (moodvie@hotmail.es), and application description (Moodvie es un recomendador de películas basado en temperamento). On the right side of the main panel, there is a list of links for managing the application: 'Editar configuración', 'Página de perfil de la aplicación', 'Estadísticas', 'Modere sus comentarios', 'Traducciones', 'Publicidad', and 'Restablecer la clave secreta de la aplicación'.

En la configuración tendremos que rellenar algunos campos, como son el nombre de la aplicación, la url del servidor que contiene la aplicación web, así como la ruta relativa dentro de Facebook que dará acceso a la aplicación (canvasurl), redirigiendo sus conexiones al servidor. En nuestro caso son: <http://supergaia.fdi.ucm.es:8060/moodvie/> y <http://apps.Facebook.com/moodvie/> respectivamente.

6.3 Casos de uso contemplados en la aplicación web

A continuación se muestra un diagrama de casos de uso con los casos contemplados para nuestra aplicación web.



6.3.1 Registrarse en la aplicación

El usuario debe poder acceder a la aplicación y ser capaz de otorgarnos permisos para poder acceder a sus datos.

6.3.2 Responder el test

El test estará accesible para que el usuario lo rellene tantas veces como quiera.

6.3.3 Introducir la provincia

Para su correcto funcionamiento, la aplicación debe permitir al usuario indicarle de qué provincia desea consultar las películas.

6.3.4 Ver las películas en cartelera

El usuario podrá consultar las películas proyectadas en su provincia.

6.3.5 Ver información sobre los cines

Al igual que se hace con la información de las películas, el usuario podrá consultar la información sobre los cines de su provincia.

6.3.6 Conocer cuándo se proyecta cada película

Dentro de la información de cada cine, además se podrá consultar los horarios de los horarios de los pases de cada película.

6.3.7 Obtener una recomendación

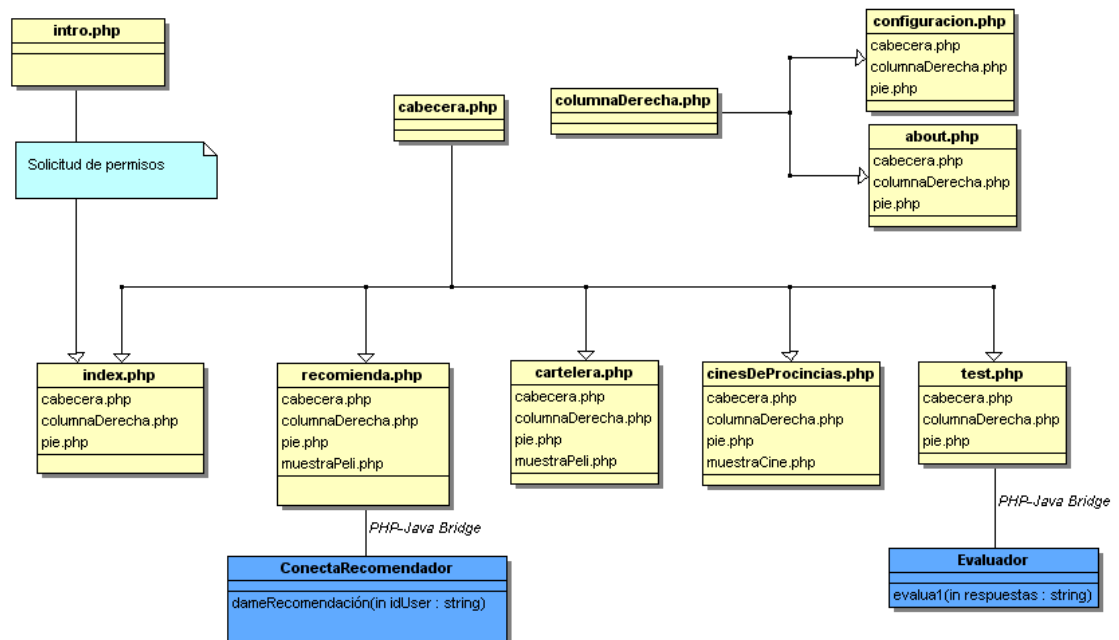
Además, el usuario podrá solicitarle al sistema que le recomiende películas siempre que este haya rellenado el test.

6.4 Desarrollo de la aplicación web

Para desarrollar una aplicación web que interaccione con Facebook se debe tener en cuenta continuamente la comunicación entre la aplicación y Facebook, asegurándonos siempre de que el usuario de nuestra aplicación lo sea también de Facebook. Esto es posible gracias al uso de la API de Facebook, que nos facilita los mecanismos necesarios para acceder a la información de cada uno de nuestros usuarios. Un paso necesario para poder tener acceso a dichos datos es que el propio usuario nos permita hacerlo; para ello, Facebook provee una interfaz transparente para el usuario, de modo que simplemente debe aceptar los permisos solicitados.

6.4.1 Estructura

A continuación se muestra la estructura de la aplicación web desarrollada:



6.4.2 Descripción de los archivos

6.4.2.1 INTRO.PHP

Contiene una breve descripción sobre la aplicación e invita al usuario a registrarse para tener acceso a ella. Solo se puede acceder a este archivo si el usuario no forma parte del sistema, en otro caso, será redirigido automáticamente a `index.php`.

6.4.2.2 INDEX.PHP

Es la página principal de la aplicación, donde se le da la bienvenida al usuario y se le muestra un pie chart con sus temperamentos. En caso de que no haya rellenado el test o la provincia, se le invita a que lo haga.

6.4.2.3 RECOMIENDA.PHP

En esta página la aplicación web se conecta a java para llamar al método *dameRecomendacion*, que devuelve las *n* películas con mejor puntuación para el usuario cuyo identificador hemos pasado como parámetro.

6.4.2.4 CARTELERA.PHP

Muestra una lista con las películas que se proyectan en la provincia del usuario.

6.4.2.5 CINESDEPROVINCIAS.PHP

Este archivo el usuario puede consultar los cines de su provincia.

6.4.2.6 TEST.PHP

Contiene el test de Keirsey para que el usuario lo pueda rellenar tantas veces como quiera. En caso de que el usuario ya tenga unos resultados guardados, se le avisará de que estos serán sobrescritos. Para obtener los resultados del test se le envían al método, implementado en java, *evalúa*.

6.4.2.7 CONFIGURACION.PHP

En esta página le usuario podrá modificar la provincia de la que se le muestre la información.

6.4.2.8 ABOUT.PHP

Donde se le muestra al usuario información acerca de la aplicación.

NOTA

Todos los documentos anteriores, salvo `Intro.php`, contienen a los archivos `cabecera.php` y `columnaDerecha.php`, que hacen posible la navegación entre ellos.

6.4.3 Problemas encontrados

6.4.3.1 TAMAÑO DE PÁGINA

Debido a problemas de comunicación entre nuestra aplicación y Facebook, Facebook no adapta bien el tamaño de la ventana al de nuestra aplicación, haciendo así que se entrecorten algunas páginas de nuestra aplicación. Como solución se ha propuesto la implementación de paginación, mostrando así menos resultados y reduciendo considerablemente el tamaño de cada página presentada al usuario.

6.4.3.2 API DE FACEBOOK

Un problema recurrente en el desarrollo de la aplicación ha sido el empleo de la API de Facebook, tanto PHP como Javascript. A pesar de facilitar la interacción con el usuario, debido a sus constantes cambios y actualizaciones se altera el comportamiento de las versiones anteriores (ya estables) de la aplicación, lo cual supone un continuo trabajo de desarrollo para solucionar nuevos comportamientos no contemplados en dichas versiones.

7 EXPERIMENTOS Y RESULTADOS

Teniendo tres sistemas de recomendación basados en el conocimiento que hemos concebido, es necesario someterlos a una serie de pruebas para comprobar su funcionamiento. Sin un estudio de los resultados obtenidos, sería imposible saber si la investigación va en buen camino.

7.1 Lista de películas

Para poder realizar las pruebas de un recomendador de películas, lo primero que necesitamos es una lista de películas. Se ha creado con este propósito, una lista de películas con su correspondiente distribución de géneros, que será usada como batería de pruebas para este estudio. Las películas que forman parte de la lista han sido seleccionadas por los miembros del grupo, intentando cubrir todos los géneros que se tienen en cuenta en las tablas de conocimiento. Cada una de ellas es representante, debido a un mayor porcentaje de género, de alguno de los once géneros que encontraremos en el sistema de recomendación. La proporción de géneros ha sido otorgada por nosotros, y no por el categorizador de géneros a partir de la sinopsis, para evitar el ruido que el categorizador pudiera meter a la hora de clasificar las películas. Vemos a continuación la lista de películas que forman la batería:

Películas				
Jumanji	70% Aventura	30% Fantasía		
La lista de Schindler	90% Drama	10% Documental		
La Guerra de las Galaxias	70% Ciencia Ficción	20% Acción	10% Aventura	
PulpFiction	40% Drama	20% Acción	20% Comedia	20% Thriller
El Exorcista	80% Terror	20% Thriller		
El Señor de los Anillos	70% Fantasía	15% Acción	15% Aventura	
Algo Pasa con Mary	85% Comedia	15% Romántica		
Casablanca	80% Romántica	20% Drama		
Un mundo de fantasía	90% Fantasía	10% Aventuras		
Psicosis	85% Thriller	10% Terror	5% Romántica	
La Jungla de Cristal	90% Acción	10% Thriller		
Bowling for Columbine	90% Documental	10% Drama		
ToyStory	60% Animación	20% Aventuras	20% Comedia	
The Ring	85% Terror	15% Thriller		

7.2 Lista ideal de recomendación

A partir de la lista de películas propuestas, los integrantes del grupo llevamos a cabo una recomendación realizada por nosotros mismos. Para ello, creamos cuatro usuarios sintéticos de la aplicación. Los usuarios representan cada uno de los cuatro temperamentos de Keirsey. Los usuarios creados son:

Usuario	% Artesano	% Idealista	% Guardián	% Racional
Mr. Artesano	85	5	5	5
Mr. Idealista	5	85	5	5
Mr. Guardián	5	5	85	5
Mr. Racional	5	5	5	85

Para cada uno de los usuarios planteados, genera cada uno de los miembros del grupo una recomendación, ordenando las catorce películas propuestas. A cada película, se le asigna una puntuación dependiendo de la posición que ocupe en la lista, la primera película recibe catorce puntos, mientras que la catorceava recibe un solo punto. Si sumamos las puntuaciones otorgadas por cada uno de los miembros del equipo, obtenemos la lista definitiva.

Esta lista es la lista ideal, ya que es el resultado obtenido a partir de tres personas que han leído y entendido en qué consiste cada uno de los temperamentos. Las listas de recomendación resultantes, en las que se ven las películas con su puntuación (a mayor puntuación, mejor posición en la lista) son las siguientes:

Películas	Mr. Artesano	Mr. Idealista	Mr. Guardián	Mr. Racional
Jumanji	14	9	1	4
La lista de Schindler	1	13	14	11
La Guerra de las Galaxias	12	6	10	13
PulpFiction	10	8	8	12
El Exorcista	7	3	7	3
El Señor de los Anillos	11	5	9	8
Algo Pasa con Mary	6	11	2	5
Casablanca	2	14	12	6
Un mundo de fantasía	5	10	5	7
Psicosis	4	4	11	10
La Jungla de Cristal	13	2	4	2
Bowling for Columbine	3	7	13	14
ToyStory	9	12	6	9
The Ring	8	1	3	1

7.2.1 Resultados de los recomendadores

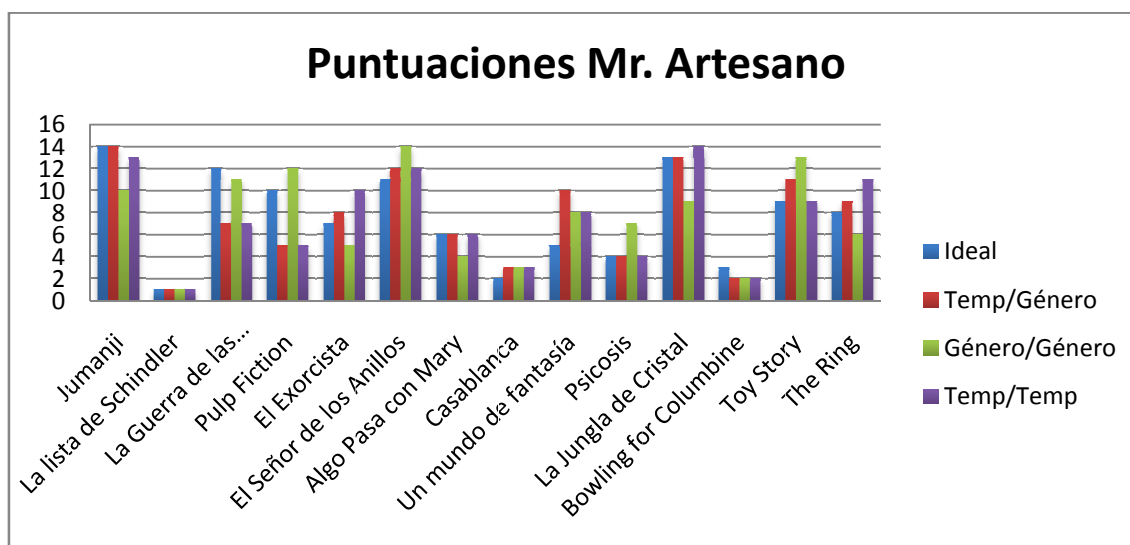
Una vez tenemos la lista de recomendación ideal, debemos compararla con los resultados obtenidos por los recomendadores, donde nuevamente las puntuaciones vendrán representadas por valores comprendidos entre el catorce y el uno. Para ello, usaremos la batería de películas creada con cada uno de los tres recomendadores (Temperamento/Género, Género/Género y Temperamento/Temperamento), para así poder comprobar cuánta similitud guardan sus resultados respecto a la lista ideal.

La estructura que se seguirá en las pruebas para los cuatro usuarios, consistirá en la lista generada por cara recomendador. Después se mostrará una gráfica donde se podrán comparar visualmente las posiciones respecto a la recomendación ideal. Luego mostraremos la diferencia acumulada de puntuaciones (cuántas posiciones ha errado el recomendador respecto a la lista ideal, por ejemplo, si la lista ideal le da a Jumanji 14 puntos y el recomendador le otorga 13, la diferencia será de dos puntos).

7.2.1.1 RESULTADOS DE PUNTUACIÓN PARA MR. ARTESANO

Películas	Temp/Género	Género/Género	Temp/Temp
Jumanji	14	10	13
La lista de Schindler	1	1	1
La Guerra de las Galaxias	7	11	7
Pulp Fiction	5	12	5
El Exorcista	8	5	10
El Señor de los Anillos	12	14	12
Algo Pasa con Mary	6	4	6
Casablanca	3	3	3
Un mundo de fantasía	10	8	8
Psicosis	4	7	4
La Jungla de Cristal	13	9	14
Bowling for Columbine	2	2	2
ToyStory	11	13	9
The Ring	9	6	11

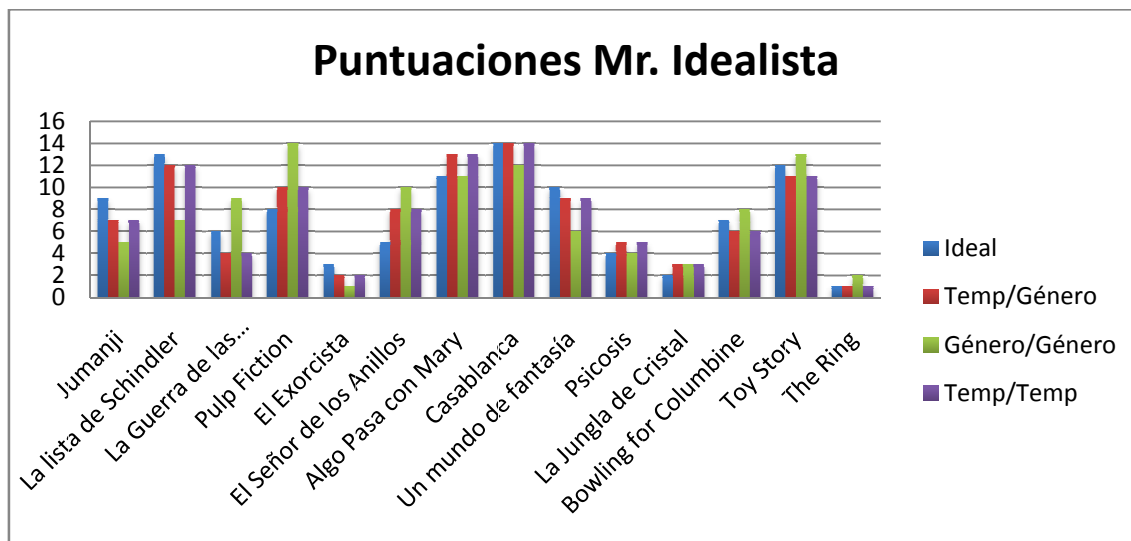
	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	22	32	24



7.2.1.2 RESULTADOS DE PUNTUACIÓN PARA MR. IDEALISTA

Películas	Temp/Género	Género/Género	Temp/Temp
Jumanji	7	5	7
La lista de Schindler	12	7	12
La Guerra de las Galaxias	4	9	4
Pulp Fiction	10	14	10
El Exorcista	2	1	2
El Señor de los Anillos	8	10	8
Algo Pasa con Mary	13	11	13
Casablanca	14	12	14
Un mundo de fantasía	9	6	9
Psicosis	5	4	5
La Jungla de Cristal	3	3	3
Bowling for Columbine	6	8	6
ToyStory	11	13	11
The Ring	1	2	1

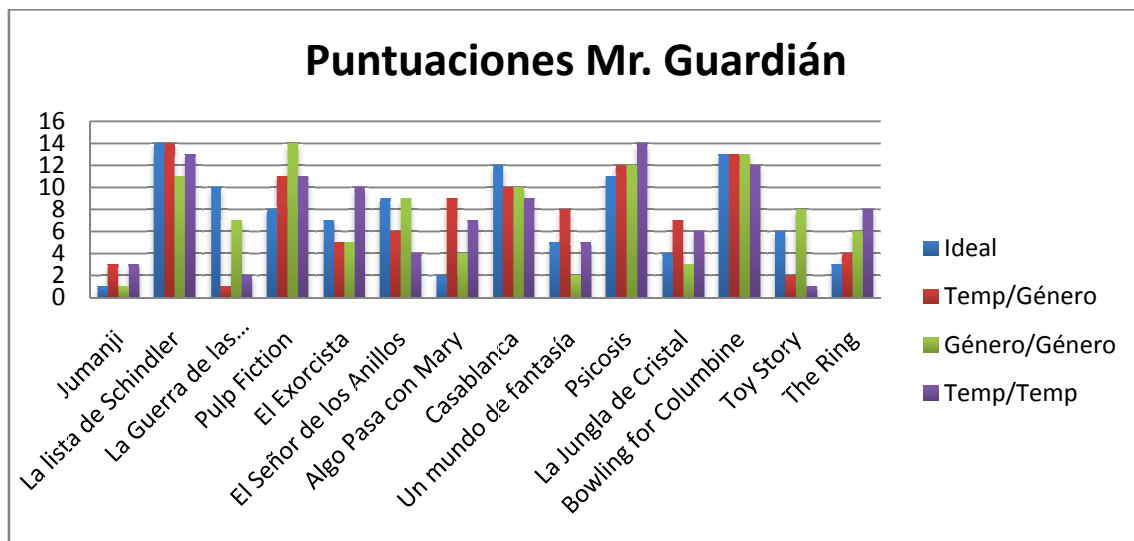
	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	18	36	18



7.2.1.3 RESULTADOS DE PUNTUACIÓN PARA MR. GUARDIÁN

Películas	Temp/Género	Género/Género	Temp/Temp
Jumanji	3	1	3
La lista de Schindler	14	11	13
La Guerra de las Galaxias	1	7	2
Pulp Fiction	11	14	11
El Exorcista	5	5	10
El Señor de los Anillos	6	9	4
Algo Pasa con Mary	9	4	7
Casablanca	10	10	9
Un mundo de fantasía	8	2	5
Psicosis	12	12	14
La Jungla de Cristal	7	3	6
Bowling for Columbine	13	13	12
ToyStory	2	8	1
The Ring	4	6	8

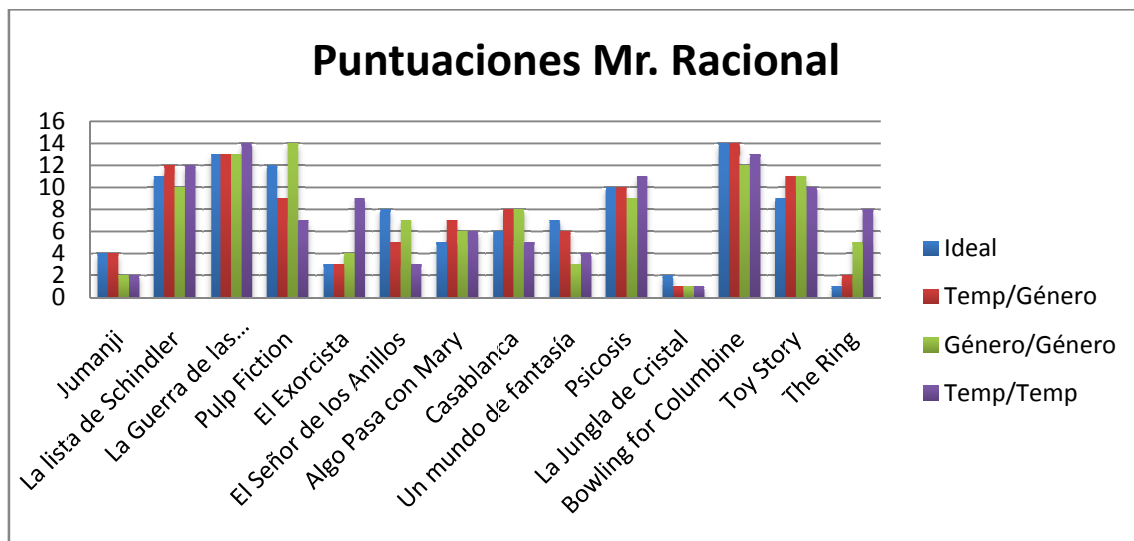
	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	40	28	46



7.2.1.4 RESULTADOS DE PUNTUACIÓN PARA MR. RACIONAL

Películas	Temp/Género	Género/Género	Temp/Temp
Jumanji	4	2	2
La lista de Schindler	12	10	12
La Guerra de las Galaxias	13	13	14
Pulp Fiction	9	14	7
El Exorcista	3	4	9
El Señor de los Anillos	5	7	3
Algo Pasa con Mary	7	6	6
Casablanca	8	8	5
Un mundo de fantasía	6	3	4
Psicosis	10	9	11
La Jungla de Cristal	1	1	1
Bowling for Columbine	14	12	13
ToyStory	11	11	10
The Ring	2	5	8

	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	16	24	36



7.2.2 Observaciones

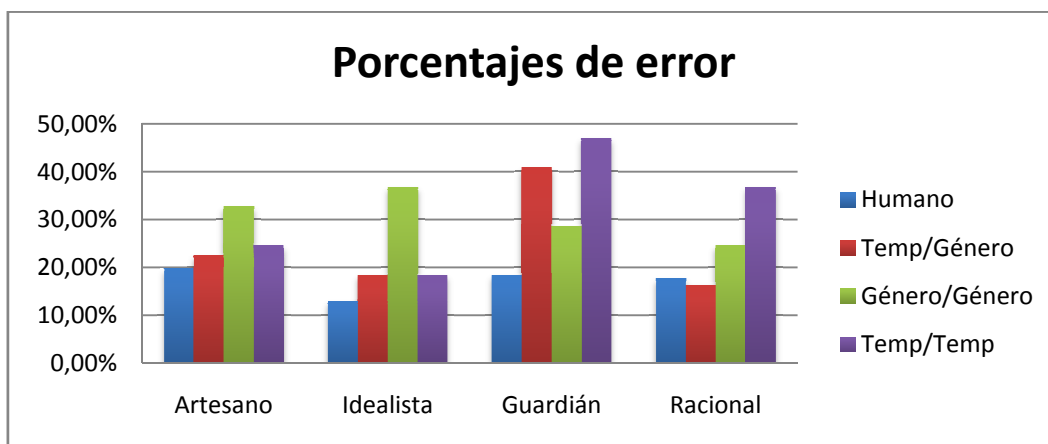
Tras el estudio realizado se puede observar cómo los recomendadores, pese a cierto error, se asemejan bastante a la recomendación ideal dada por los componentes del grupo. Dependiendo del temperamento que se esté analizando, cada recomendador tiene un nivel de acierto distinto. Puede verse por ejemplo cómo en el temperamento Guardián se obtienen buenos resultados con el recomendador Género/Género, mientras que con los otros recomendadores se obtienen malos resultados. En cambio, para otro temperamento, como puede ser el idealista sucede a la inversa, el recomendador Género/Género obtiene mucho error, mientras que los otros dos recomendadores obtienen resultados muy acertados.

Podemos mostrar, a modo comparativo, el error medio que habíamos obtenido los miembros del proyecto respecto a la lista ideal resultante de cada temperamento:

	Artesano	Idealista	Guardián	Racional
Diferencia Puntuaciones	19,3333333	12,6666667	18	17,3333333

Vemos que el error humano medio es bastante alto, teniendo en cuenta que la lista ideal es una media de nuestros resultados por separado. Como vemos, los resultados de error humano no distan tanto de los obtenidos por los recomendadores.

Si tenemos en cuenta, que el error máximo posible tiene un valor de 98 (resultado de una lista completamente invertida respecto a la lista ideal), podemos calcular los porcentajes de error de todos los recomendadores, por cada temperamento.



Se observa cómo para todos los temperamentos, existe un recomendador que obtiene mejores resultados de los demás, siendo estos similares al error medio humano (o incluso mejorándolo en algún caso). Además, como se explica en el apartado de “Recomendación Unificada”, si unimos los tres recomendadores a la vez, podemos obtener mejores resultados que cuando trabajan por separado.

8 RECOMENDACIÓN UNIFICADA

La recomendación unificada es el resultado de la unión de los tres recomendadores por separado. Hasta ahora, se habían descrito en el apartado de “Sistema Recomendador” los tres recomendadores que se habían implementado independientemente (el recomendador Temperamento/Género, el Género/Género y el Temperamento/Temperamento), mientras que en esta sección se explicará su funcionamiento colaborativo.

Sobre cada uno de los tres recomendadores independientes se había realizado un estudio tal y como se puede observar en el apartado de “Análisis de pruebas”. De ese estudio se concluía que los tres recomendadores funcionaban bien, pero no igual. Casos en los que un recomendador nos ofrecía buenos resultados únicamente para un temperamento concreto, otros en los que se ofrecían resultados peores (que no malos) pero con regularidad en todos los temperamentos, etc.

La intención de la unificación de recomendadores es la de reforzar los aciertos demostrados por ellos a la vez que se mitigan sus errores. Si los tres recomendadores ofrecen su lista resultante y se realizan una serie de operaciones para combinarlas, podemos llegar a una lista de recomendación final de mayor fiabilidad que las que se ofrecían por separado.

8.1 *Proceso de unificación*

La idea del proceso de unificación consiste en asignar un porcentaje de peso a cada lista ofrecida por el recomendador, en la que las puntuaciones de las películas se multiplicarán por el peso característico del recomendador que las ha generado, sumar las puntuaciones de películas que coincidan en distintos recomendadores, y devolver una nueva lista con los resultados globales.

Existen distintas maneras de plantear la repartición de pesos entre los recomendadores dependiendo del temperamento. Una posibilidad es la de que por cada temperamento se asigne un porcentaje mayor, otro intermedio y otro menor a cada uno de los recomendadores, siempre usando los mismos valores y basándonos en los resultados obtenidos en el “Análisis de pruebas” (por ejemplo: para el mejor recomendador un peso del 50%, para el segundo 30% y para el tercero 20%, manteniendo estos valores en todos los recomendadores). Este sistema podría ser válido, pero como se han visto en nuestras pruebas, los sistemas de recomendación no tienen comportamientos similares en todos los temperamentos, con lo cual sería en cierto modo un posible error medirlos siguiendo el mismo patrón.

Otra posibilidad es la de experimentar con distintos pesos en los recomendadores, usando la batería de pruebas de películas empleada en la fase de análisis, hasta encontrar una combinación de proporciones que minimice el error de cada temperamento. Este sistema ha sido estudiado, implementado y puesto a prueba. Se obtienen buenos resultados, pero con ligeros problemas conceptuales. Si la batería de pruebas fuese de un tamaño mucho mayor que la usada, se podrían extraer realmente buenas conclusiones, y seguramente esta opción sería muy válida. El problema reside en que en la batería, al ser de un tamaño pequeño, el hecho de cambiar los pesos asignados a cada uno de los recomendadores con la intención de encontrar unos porcentajes que minimicen el error, puede ser inválido para otra batería de películas distintas. Podríamos tener mínimos de error locales para baterías diferentes, pero que al cambiar las películas, perdiesen su validez. Esto es debido a que las pequeñas modificaciones de porcentajes servirían para corregir problemas del caso estudiado en ese preciso momento y los errores se anularían entre ellos por mera coincidencia para ese caso concreto.

La última posibilidad planteada es la de distribuir los pesos a partir del número de errores obtenidos en las pruebas. Si sumamos el número de aciertos (errores máximos posibles, que tiene un valor de 98, menos errores cometidos) de cada recomendador por cada temperamento y hallamos el número máximo de aciertos, podemos sacar el porcentaje de aciertos relativo dado por cada recomendador, obteniendo así el peso que deberá tener cada uno de ellos. De este modo, aunque volvamos a basarnos en una batería pequeña, evitamos la distribución de pesos producida por pura coincidencia, basándonos esta vez en el porcentaje real que ha habido de acierto en cada uno de los recomendadores.

De este modo, los porcentajes por temperamento quedarían tal y como se muestran a continuación:

8.1.1.1 TEMPERAMENTO ARTESANO			
	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones (error)	22	32	24
Acierto (error máximo - error)	76	66	74
Peso Recomendador (acierto / \sum aciertos)	35%	31%	34%

8.1.1.2 TEMPERAMENTO IDEALISTA			
	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	18	36	18
Acierto	80	62	80
Peso Recomendador	36%	28%	36%

8.1.1.3 TEMPERAMENTO GUARDIÁN

	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	40	28	46
Acierto	58	70	52
Peso Recomendador	32%	39%	29%

8.1.1.4 TEMPERAMENTO RACIONAL

	Temp/Género	Género/Género	Temp/Temp
Diferencia Puntuaciones	16	24	36
Acierto	82	74	62
Peso Recomendador	38%	34%	28%

En las tablas mostradas encontramos los pesos que se asignan a cada recomendador por cada uno de los cuatro temperamentos de Keirsey. Estos pesos se encuentran en el código del proyecto en la clase “RepartoRecomendadores”. De este modo se pueden realizar pruebas modificando los pesos de un modo muy sencillo.

8.1.2 Pesos de unificación individualizados

Una vez obtenidos los pesos asociados a cada recomendador por cada temperamento, el proceso de individualización por la distribución de pesos del usuario es el resultado de unas simples operaciones aritméticas.

Como ya se ha explicado anteriormente en la memoria, cada usuario dispone de una distribución temperamentos. La única tarea que se debería llevar a cabo para encontrar el peso de un recomendador individualizado para un usuario concreto sería la de obtener por cada recomendador, la suma de pesos que tiene para cada temperamento al multiplicarse por el porcentaje de temperamento de ese usuario. Se puede observar con detalle en la siguiente fórmula:

$$\forall \text{Recomendador } r,$$

$$Peso_r = \sum_{t=\text{Artesano}}^{\text{Racional}} \text{tablaRepartoRecomendadores}(r, t) * \text{porcentaje}(\text{Usuario}, t)$$

8.1.3 Recomendación unificada individualizada

Una vez tenemos los pesos exactos que deben aplicarse a los recomendadores de un usuario concreto gracias a los cálculos anteriores, podemos pasar a la fase de recomendación.

El proceso consiste en, por separado, calcular la lista de recomendación de un usuario con cada uno de los tres recomendadores. De este modo, obtenemos tres listas distintas de películas. A cada una de las películas de esa lista, se le atribuye una puntuación en base al orden obtenido en la recomendación (la primera obtendrá diez puntos mientras que a la última se le otorgará un único punto).

Una vez se han puntuado las listas con esta escala, pasarán a multiplicarse sus puntuaciones por el valor obtenido en los pesos individualizados de los recomendadores. De este modo, la puntuación obtenida se ve modificada en base a la importancia que el recomendador en cuestión tenga para dicho usuario.

Si la película aparece en varios recomendadores, sus puntuaciones se irán sumando. En caso de que solamente aparezcan en un recomendador, su puntuación será la calculada con la operación que se ha explicado en el párrafo previo.

Tras obtener la evaluación de todas las películas se procede a la extracción de las X películas con más puntuación, que pasará a ser la lista definitiva de películas recomendadas. Con este sistema conseguimos el reforzando de las recomendaciones que se podían realizar por un recomendador aislado y obteniendo resultados de recomendación dirigidos hacia un individuo concreto usuario de la aplicación.

9 RECOMENDADOR CON ESTADOS DE ÁNIMO

Un estado de ánimo es un estado emocional en el que puede encontrarse una persona durante un periodo más o menos largo. Difieren de las emociones, ya que los estados de ánimo son más duraderos y se sienten con menos intensidad. Una persona tiene un temperamento, que es permanente, pero dependiendo de su estado de ánimo puede preferir unas cosas u otras.

Por este motivo, se debe considerar la posibilidad de un recomendador no sólo basado en temperamentos, sino en el que el usuario tenga la posibilidad de especificar estados de ánimo para orientar la búsqueda hacia determinados resultados.

9.1 Ideas propuestas

La primera idea es la de que el usuario especifique su propio estado de ánimo, y así se pueda realizar una recomendación teniendo esta información en cuenta. El problema de este método es la elección de películas acertadas sobre las que realizar la recomendación. Por ejemplo, si un usuario se encuentra sumido en la tristeza, elegir qué película recomendarle no es tarea fácil. Se podría recomendar una triste, por afinidad a su estado. También se le podría recomendar una comedia, para contrarrestar su estado y hacerle feliz. Por otro lado, es posible que una película sorprendente le haga olvidar sus pesares. Motivos como estos son los que nos hacen descartar la idea de recomendar en base al estado de ánimo que envuelve al usuario en un momento determinado.

La alternativa encontrada, que elimina los problemas encontrados en la idea previa, y al mismo tiempo dota al usuario de un grado mayor de libertad, es la de que el usuario pueda elegir un objetivo anímico. La posibilidad de que la persona determine cómo quiere sentirse durante y tras el visionado de la película. De este modo, se puede realizar otro tipo de recomendación dirigida hacia un fin concreto. Los estados de ánimo objetivo que se han tenido en cuenta en nuestra aplicación son los siguientes:

- Sorprendido
- Alegre
- Asustado
- Triste
- Relajado o tranquilo

Cada uno de los estados de ánimo planteados podrá ser seleccionado por el solicitante de la recomendación a modo de objetivo. De este modo, la recomendación se orientará hacia determinadas películas.

9.1.1 Proceso de recomendación con estado de ánimo objetivo

Una de las intenciones planteadas para la recomendación fue la de modificar las puntuaciones obtenidas por el recomendador temperamental. De este modo, se podía añadir o restar puntuación a la película, teniendo en cuenta el estado objetivo del usuario y la distribución de géneros de la película. Es decir, una película con una puntuación determinada X , podría resultar con puntuación $X+Y$, siendo Y el valor obtenido al operar con una variable multiplicativa por el porcentaje de género orientado al estado de ánimo deseado.

A modo de ejemplo, supongamos que el usuario quiere estar alegre. La puntuación para el usuario de una película determinada, con 30% de comedia, es de 80. Con una variable multiplicativa de 10, la puntuación de la película sería $80 + (10 * 0.3) = 83$.

De este modo, cuanto más peso tenga un género relacionado con el deseo anímico del usuario, más aumentaba la puntuación. Del mismo modo, podríamos clasificar géneros contrarios al deseo del usuario que bajaran la puntuación en vez de incrementarla.

El problema de realizar estas operaciones es que se altera el resultado obtenido por los cálculos de la tabla de temperamentos. Se pueden dar casos en los que una película que no guste demasiado, pero con un alto porcentaje de un género determinado –el asociado al estado de ánimo buscado– gane a una que el usuario disfrutaría mucho más y tenga un porcentaje significativamente menor del género antes mencionado. También cabe la posibilidad de que una película, pese a no tener absolutamente nada de un género relacionado con el estado objetivo, fuese la que lograra una mayor puntuación, debido a que obtiene la puntuación más alta incluso aplicando modificadores a las demás películas. Por estos motivos, este tipo de recomendación fue aparcada y se buscó otra propuesta para solventar el problema.

La alternativa que elimina el obstáculo planteado es la de emplear un filtro de películas. Del otro modo, el pilar de nuestro recomendador podía verse debilitado por el efecto de multiplicadores que desvirtuaban los resultados. Ahora, en cambio, las películas que cumplan determinadas condiciones para satisfacer a un usuario que busca un estado de ánimo concreto son las seleccionadas para el proceso de recomendación, descartando aquellas que no deberían recomendarse. El sistema implementado consta de una variable límite que ajusta el ancho del filtro determinando qué películas pasan a la fase de recomendación y cuáles obtienen directamente una puntuación nula.

Se han relacionado los géneros filmicos con los estados de ánimo a la hora de establecer el filtro, con una constante X de porcentaje de género:

- Sorprendido: Aquellas películas que sobrepasen el X% de Acción o el X% de Aventura.
- Alegre: Aquellas películas que sobrepasen el X% de Comedia.
- Asustado: Aquellas películas que sobrepasen el X% de Terror o el X% de Thriller, siempre y cuando la Animación esté por debajo del X%.
- Triste: Aquellas películas que sobrepasen el X% de Drama.
- Relajado o tranquilo: Aquellas películas que sobrepasen el X% de Romántica o el X% de Documental.

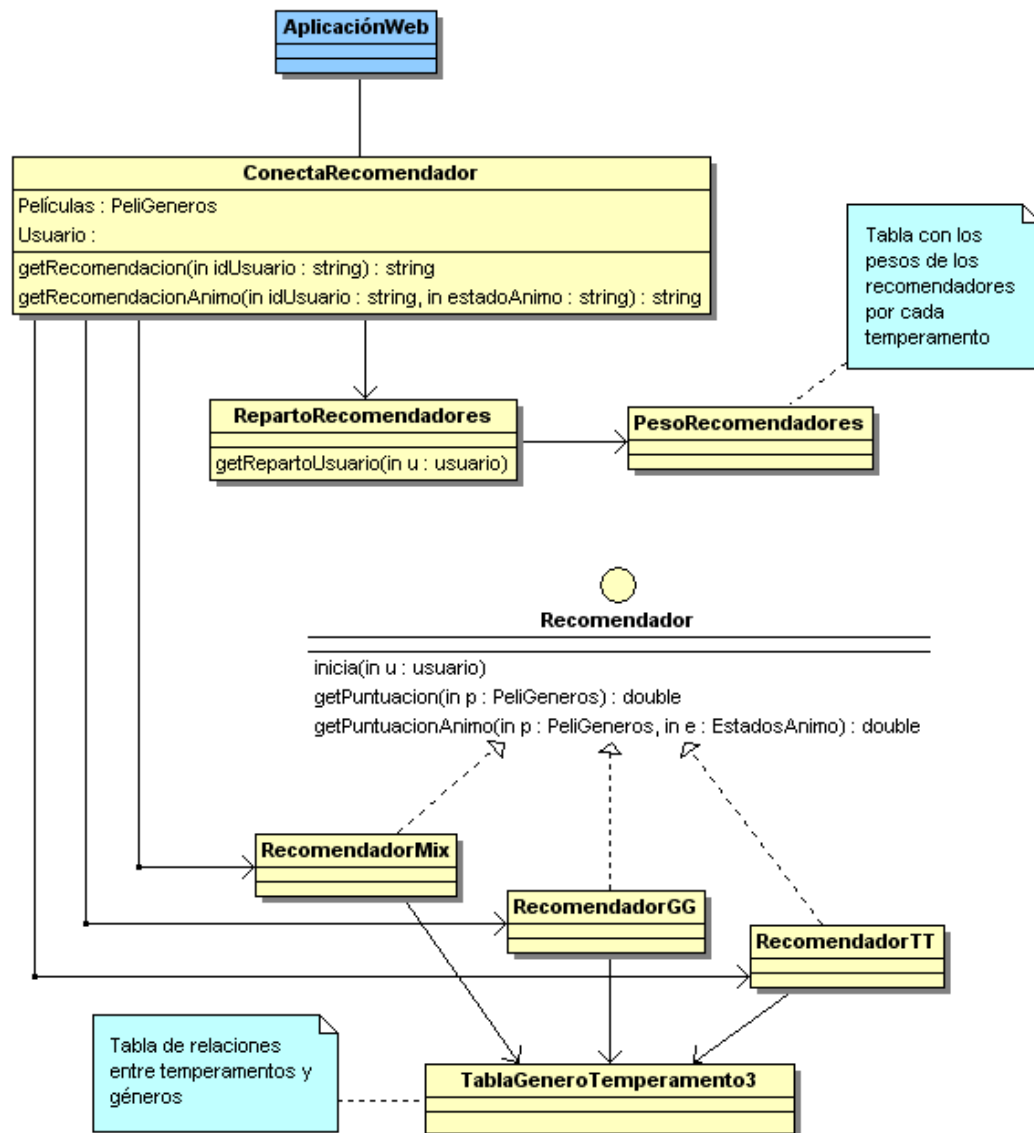
Con este nuevo sistema no se ve alterado el sistema propio del recomendador. La recomendación se basa en el conocimiento extraído de los temperamentos de Keirsey, pero esta vez, la puntuación solamente se calcula sobre las películas que superen el filtro previo.

La aplicación, de este modo, cubre las necesidades de los usuarios. Ahora el usuario puede elegir una recomendación general, basándose en sus temperamentos, o una recomendación que sigue basándose en los temperamentos pero cuyo resultado se encuentra entre las películas que pasen el filtro determinado por su estado de ánimo objetivo elegido.

10 DIAGRAMAS DEL RECOMENDADOR

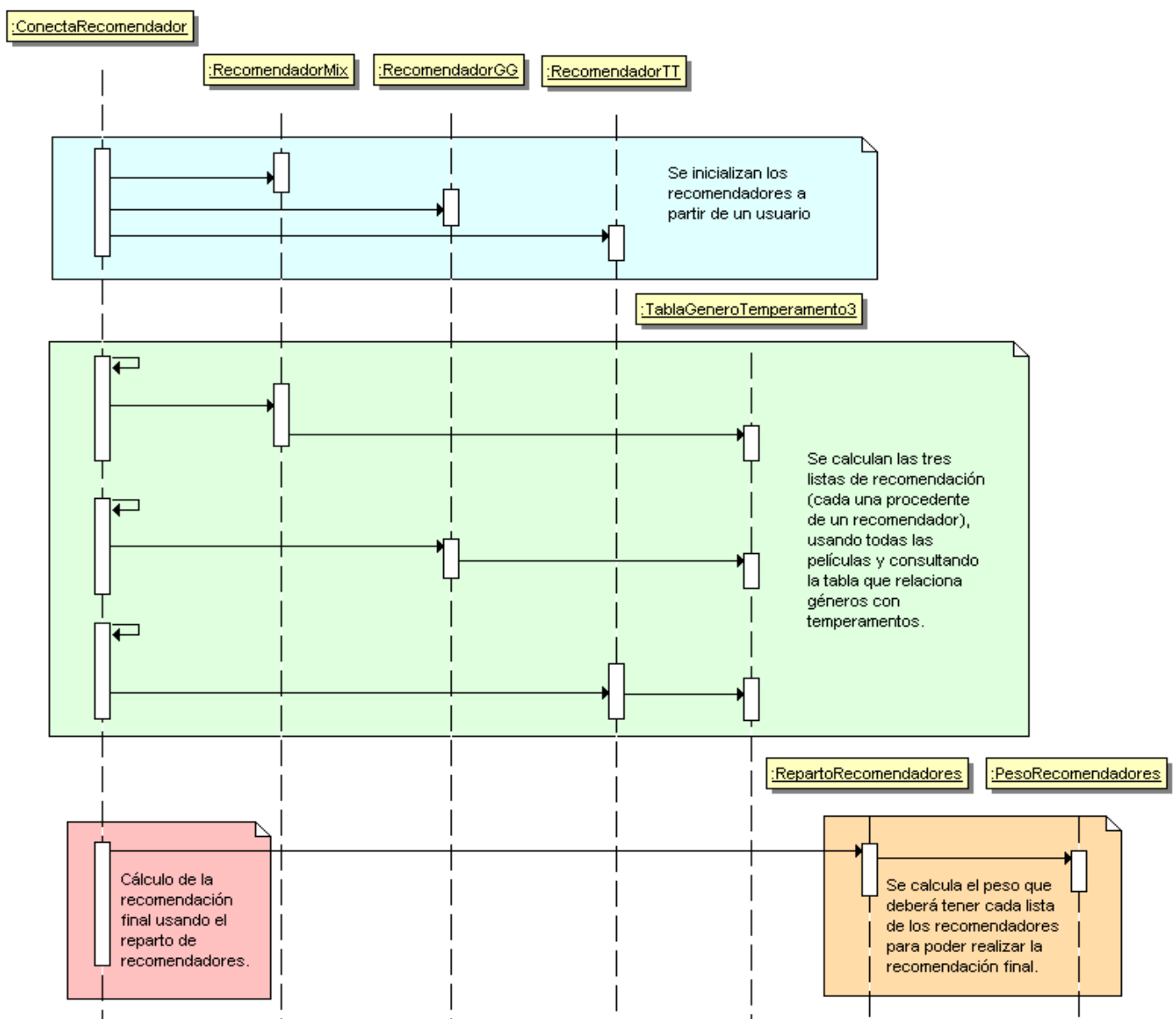
En este apartado se muestran a modo de resumen esquemático una serie de diagramas para ayudar a entender conceptualmente la implementación.

10.1.1 Diagrama de clases



La clase **ConectaRecomendador** es la encargada de gestionar el proceso de recomendación en sí. A partir del usuario de la aplicación web, se extraen todas las películas que están siendo proyectadas en su provincia, con sus correspondientes géneros. Los tres recomendadores (**RecomendadorMix**, **RecomendadorGG**, **RecomendadorTT**) implementan la interfaz **Recomendador**, que tiene un método de inicialización del recomendador a partir del usuario, y métodos de obtención de puntuaciones generales y con estado de ánimo. Los tres recomendadores acceden a la **TablaTemperamento3** para extraer la información necesaria para obtener las puntuaciones. Por otro lado, la clase **RepartoRecomendadores** se encarga de calcular los pesos que deberá tener cada recomendador para un usuario concreto consultando la tabla de pesos de recomendadores **PesoRecomendadores**, que ha sido explicada en el apartado de “Unificación de recomendadores”.

10.1.2 Diagrama de secuencia



En el diagrama de secuencia se puede seguir el orden en que se llaman a las clases para llevar a cabo la recomendación. Por un lado, la primera tarea del conector es la de inicializar los tres recomendadores. Después, hay un bucle por recomendador encargado de calcular las puntuaciones de todas las películas y generar las listas de recomendación. Por último, encontramos el cálculo final de la recomendación, que se sirve de los pesos calculados por el reparto de recomendadores. De este modo tenemos los resultados finales listos para ser devueltos a la aplicación web.

11 FEEDBACK DE LOS USUARIOS

Para evaluar nuestra aplicación, hemos solicitado a los usuarios que la han utilizado que rellenen un sencillo formulario que nos dé una ligera idea sobre su punto de vista sobre diversos aspectos de nuestra aplicación. En las siguientes secciones se mostrará el formulario y los resultados obtenidos.

11.1 Formulario

Pregunta	Puntuación (1..5)
1. Evaluación de la aplicación	
1.1. Valora de 0 a 5 lo fácil que te ha resultado entender y utilizar la aplicación (0=nada entendible, 5=muy entendible)	
1.2. Valora de 0 a 5 el diseño de la aplicación (0=mal diseñada, 5=muy bien diseñada)	
1.3. Valora de 0 a 5 la claridad del test (0=poco claro, 5=muy claro)	
1.4. Valora de 0 a 5 tu satisfacción de uso de la aplicación (0=poco satisfecho, 5=muy satisfecho)	
2. Evaluación de la recomendación	
2.1. Valora de 0 a 5 el número de películas de la recomendación que conoces (0= cero películas, 5=cinco películas)	
2.2. Evalúa de 0 a 5 tu gusto sobre la primera película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	
2.3. Evalúa de 0 a 5 tu gusto sobre la segunda película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	
2.4. Evalúa de 0 a 5 tu gusto sobre la tercera película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	
2.5. Evalúa de 0 a 5 tu gusto sobre la cuarta película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	
2.6. Evalúa de 0 a 5 tu gusto sobre la quinta película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	
2.7. Evalúa de 0 a 5 la precisión del resultado de la recomendación (0=no creo que me gustase ninguna película, 5=creo que las 5 películas me gustarían)	

11.2 Resultados obtenidos

Los resultados aquí presentados se han obtenido mediante la media aritmética sobre las puntuaciones de cada pregunta. Esta se hace sobre un total de 17 encuestas contestadas.

Pregunta	Suma de puntuaciones	Número de respuestas	Puntuación media (0..5)
1. Evaluación de la aplicación			
1.1. Valora de 0 a 5 lo fácil que te ha resultado entender y utilizar la aplicación (0=nada entendible, 5=muy entendible)	66	17	3.88
1.2. Valora de 0 a 5 el diseño de la aplicación (0=mal diseñada, 5=muy bien diseñada)	63	17	3.70
1.3. Valora de 0 a 5 la claridad del test (0=poco claro, 5=muy claro)	60	17	3.52
1.4. Valora de 0 a 5 tu satisfacción de uso de la aplicación (0=poco satisfecho, 5muy satisfecho)	64	17	3.76
2. Evaluación de la recomendación			
3.1. Valora de 0 a 5 el número de películas de la recomendación que conoces (0= cero películas, 5=cinco películas)	46	17	2.70
3.2. Evalúa de 0 a 5 tu gusto sobre la primera película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	38	17	2.23
3.3. Evalúa de 0 a 5 tu gusto sobre la segunda película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	39	17	2.29
3.4. Evalúa de 0 a 5 tu gusto sobre la tercera película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	45	17	2.64
3.5. Evalúa de 0 a 5 tu gusto sobre la cuarta película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	41	17	2.41
3.6. Evalúa de 0 a 5 tu gusto sobre la quinta película de la recomendación (0=creo que no me gustaría nada esa película, 5=creo que me gustaría mucho esa película)	43	17	2.72
3.7. Evalúa de 0 a 5 la precisión del resultado de la recomendación (0=no creo que me gustase ninguna película, 5=creo que las 5 películas me gustarían)	49	17	2.88

11.1 Sugerencias

Además de responder el formulario, los usuarios de la aplicación tuvieron disponible un campo de sugerencias de mejora de la aplicación. A continuación enumeramos aquellas que, bajo nuestro punto de vista, son más destacadas o recurrentes por parte de los usuarios, otorgando así ideas de trabajo futuro. Otras, que no enumeraremos aquí, fueron implementadas por nosotros mismos debido a que la dificultad que presentaba su implantación era aceptable respecto a la mejora obtenida, y no suponían grandes cambios en el enfoque de recomendación propuesto.

- I. La sensación al probarla no es clara. Se obtiene una sensación de no saber si te ha recomendado esa película por tu temperamento o porque de las que hay en el cine al fin y al cabo son las mejores., especialmente en la recomendación por estado de ánimo.
- II. Introducir enlaces a trailers.
- III. Nueva variante en la recomendación, basada en la época del año, suponiendo que las películas con trama argumental más compleja no son adecuadas para, por ejemplo, el verano.
- IV. Nueva variante en la recomendación, presentando al usuario películas parecidas a una seleccionada (y que no necesariamente estén en cartelera en ese momento).
- V. Votación de películas.

La sugerencia I es un problema recurrente en los sistemas recomendadores, y no hay definido un método de actuación para solventarlo. Se barajó la opción de presentarle al usuario las puntuaciones de las películas recomendadas, así como las que seleccionaba cada recomendador por separado y cómo se unían. Sin embargo, a pesar de ser técnicamente implementable y de una dificultad no muy elevada, comprobamos la dificultad de presentar toda esa información al usuario de una manera sintética, de modo que fuese fácilmente comprensible. Ante estas dificultades decidimos no abordar el problema, pues la solución encontrada confundía más de lo que aclaraba.

La sugerencia II es recurrente, pero desafortunadamente dicha información no aparece en La Guía del Ocio. Se barajó crear una extensión para el extractor que completase la información con dichos enlaces, pero no encontramos una fuente de conocimiento que nos los asegurase. Por ello, no se implementó ninguna mejora en este aspecto.

Las sugerencias III y IV, a pesar de resultar interesantes, fueron descartadas, bien por la escasez de tiempo para afrontar el problema desde otro punto de vista totalmente diferente, o bien por ser necesario una gran base de datos de película, con la que no contábamos. Como ya se ha explicado en numerosas ocasiones, la novedad del enfoque propuesto es evitar el arranque en frío o el conocimiento previo del medio por parte del usuario.

La sugerencia V suscitó un debate importante dentro del grupo. Si bien la implementación de la votación no suponía gran trabajo, consideramos que una vez abordado este asunto, sería importante establecer otras opciones como “conocer personas con mis gustos” o, por supuesto, mejorar las recomendaciones en base a las votaciones de cada película. Este trabajo estaba claramente fuera de lo abordable en un proyecto de esta envergadura, por lo que decidimos no implementarlo. En cualquier caso, consideramos este nuevo enfoque muy interesante, pues sería compatible con la visión que hemos proporcionado, que evita el problema del arranque en frío, y mejoraría la recomendación según los gustos del usuario. De este modo, además de evitar el problema del arranque en frío, las recomendaciones podrían mejorarse mediante comparaciones sociales, es decir, “películas que le hayan gustado a un usuario con un temperamento parecido al mío”, y con comparaciones individuales –esto es, “películas con una distribución de géneros parecida a alguna de las que me haya gustado”.

APENDICE. TABLA DE TEMPERAMENTOS

TEMPERAMENTO	Artesano	Guardián	Idealista	Racional
Autoimagen: cómo se sienten y los ven los otros				
Autoestima	Artísticos	Confiabiles	Empáticos	Ingeniosos
Auto-respeto	Audaces	Caritativos	Benévolos	Autónomos
Autoconfianza	Adaptables	Respetables	Auténticos	Resueltos
Valoran				
Estado de ánimo preferido	Emocionado (exteriormente). - Buscan la emoción a través de la acción aunque no lo muestren y olvidan el cansancio o el dolor. - Animar a los demás, transmiten energía. - Se aburren fácilmente.	Interesado por temas serios: hogar, trabajo, deberes, temas sociales. - Preocupados concienzudamente por sus responsabilidades grandes y pequeñas.	Entusiasta (interiormente). - Se entusiasman con sus ideas o planes. - Muy emotivos positivos cuando hablan.	Tranquilo (incluso en situaciones caóticas). - Ocultarán los otros estados de ánimo.
Confían en	Impulso (actuando). - Siguen su instinto, son impulsivos y espontáneos, no pueden esperar. - Viven el momento, el presente.	Autoridad. - Acatan leyes, órdenes y normas sin cuestionar.	Intuición (actuando).	Razón, lógica. - A veces en la intuición. - Desconfían de la Autoridad.
Anhelan diariamente	Impacto. (en los demás) - Acciones que causen impacto, sensación, un efecto visible. - Sentirse poderosos	Pertenencia a algún grupo social. - Que puedan confirmar diariamente que son miembros importantes	Romance. - Relación idealizada (con los demás) - Las posibilidades	Logros. - Normalmente muy técnicos - Se exigen mucho - Tienen sensación de que no es suficiente lo que han conseguido, incluso que están al borde del fracaso.
Buscan en cada instante	Estímulos -Sensaciones a través de los cinco sentidos. -Vivir “todo”, probar todo, -La variedad, lo nuevo	Seguridad -Son conscientes de los peligros	Su propia Identidad -El auténtico Ser: propio y ajeno, no una máscara	Conocimiento Pragmático -Saber la condiciones necesarias y suficientes para que ocurra algo -Aprender a predecir y controlar eventos -Poner a prueba sus conocimientos con situaciones nuevas. -Buscador de conocimiento
Valoran, se sienten apreciados si hay	Trato generoso. - Disfrutan dar y recibir espontáneamente. No por obligación sino por diversión.	Gratitud. -Los demás les expresan gratitud por lo que han hecho - Sólo a los que se lo merecen.	Ser reconocidos como únicos. - Que los demás conozcan su mundo.	Atención por sus logros. - Que les pidan que expliquen sus logros y sus motivos.
Aspiran, sueñan con ser	Virtuosos. -Maestro en el manejo de una técnica, método, habilidad, instrumento o herramienta.	Ejecutivos. - Dirigir instituciones. - Ejercer la autoridad.	Sabios. - Visión filosófica de la vida. - Más allá de preocupaciones terrenas y temporales	Magos Tecnológicos. - Genio Científico.

<i>Papel social</i>				
Pareja ideal	Compañero de Juegos. -Buscan alegrar a su pareja dando alegría y placer. -Que genere diversión y emoción, no aburrimiento.	Asistente. - Que ayude en todo para construir familia cómoda y estable. - Son leales incluso en los malos tiempos.	Alma Gemela. - Que conozca sus sentimientos sin decirlos. - Le digan que son especiales. - Relación profunda y significativa.	Mente Afin. - Que compartan lo que piensan hasta que quede claro, no importa si no están de acuerdo. - Poder compartir temas abstractos. - Eligen pareja haciendo un estudio empírico cuidadoso: una pareja debe durar toda la vida.
Rol paterno	Liberador - Dominan los límites respetando a los demás. - Muestran iniciativa, aventura y a manejar muchas técnicas y herramientas. - Animar a irse pronto de casa.	Socializador, cortés. - Buscan que sus hijos estén de acuerdo con la comunidad, sean civilizados con cultura: sean útiles	Armonizador. - Crean "autoimagen" positiva y sana.	Individuador.
Liderazgo	Negociador. - Descubren lo que puede ser una ventaja cuando negocian. - Improvisan con lo que tienen a mano, buscan ventajas para llegar a acuerdos. - Oportunistas	Estabilizador. - Estableciendo normas muy bien estudiadas sobre qué, cómo y quién hace las cosas. - Vigilando que todos sigan esas normas.	Catalizador. - Facilitador, motivador o acelerador. - Buenos sentimientos en los subordinados. - Cooperación	Visionario.

I ANEXO I. DOCUMENTACIÓN DEL SISTEMA IMPLEMENTADO

I.1 Manual del Usuario

En este apartado se explicarán las funcionalidades de Moodvie para los distintos tipos de usuario.

I.1.1 Usuario común

El usuario común, aquel al que va dirigida la aplicación, únicamente tiene acceso al front-end del sistema recomendador, esto es, la aplicación en Facebook. Desde allí podrá:

1. Registrarse en la aplicación
2. Responder el test
3. Introducir la provincia
4. Ver las películas en cartelera
5. Ver información sobre los cines
6. Conocer cuándo se proyecta cada película.
7. Obtener una recomendación

REGISTRARSE EN LA APLICACIÓN

La primera vez que el usuario entre en la aplicación, visualizará información similar a la siguiente imagen:



Allí se le explica al usuario, de una manera somera, qué es Moodvie, cómo funciona, y qué funcionalidad tiene. Una vez pulse en el botón Entrar en Moodvie, se le solicitarán los permisos que la aplicación requiere, esto es, publicación en el muro y acceso a su información básica de perfil. Si el usuario no estuviese identificado en Facebook, se le solicitará primero la identificación y a continuación la aceptación de los permisos necesarios para el funcionamiento de la aplicación.

I.1.1.1 RESPONDER EL TEST

Para generarse el temperamento, el usuario debe responder un test de 16 preguntas. Esto lo puede realizar desde el apartado “Realizar test”, que se encuentra en el menú principal de la aplicación. Una vez contestadas las 16 preguntas, al pulsar el botón “Enviar”, Moodvie generará una personalidad única al usuario, y volverá a la pantalla de inicio. Ahora, el usuario verá una tabla, así como un gráfico con forma de tarta, con su porcentaje de temperamentos, tanto en forma escrita como visual. De este modo, se garantiza una correcta accesibilidad de la aplicación, así como un aspecto gráfico más vistoso.

Aunque ya se haya hecho el test, siempre es posible volver a realizarlo. El proceso a seguir es el mismo que el realizado la primera vez, pero Moodvie avisará de que ya hay un temperamento generado. En caso de seguir, se eliminará el temperamento anterior, y se sustituirá por el nuevo.

I.1.1.2 INTRODUCIR LA PROVINCIA

La elección de provincia se hace desde el apartado “Configuración”, que se encuentra en el menú secundario. Allí aparece una lista desplegable con todas las provincias con las que trabaja Moodvie. Una vez seleccionada la provincia, basta con apretar el botón de “Actualizar” para que la selección se guarde en la configuración del usuario.

I.1.1.1.3 VER LAS PELÍCULAS EN CARTELERA

Una vez seleccionada la provincia, el usuario puede ver las películas que se están proyectando en dicha provincia. Para ello, basta con acceder a la opción “**Ver cartelera**” del menú principal. Desde allí se puede ver un resumen corto de cada una de las películas, la categorización de géneros asignados por el sistema y acceder a la información completa o a un listado de los cines en que se proyecta –con sus correspondientes sesiones.

facebook Buscar Inicio Perfil Buscar amigos Cuenta

Inicio Obtener recomendación Ver cartelera Ver cines Realizar test

CARTELERA

A continuación puedes ver todas las películas de tu provincia disponibles en nuestra base de datos

Prev 1 2 3 4 5 6 7 8 9 10 11 Next

Agua para elefantes

Género: Drama
Duración: 02:02:00
Calificación: Mayores de 12 años
Nacionalidad: EE.UU
Fecha de estreno: 2011

Géneros calculados:

8.3954%	9.9745%	26.1302%	20.4279%	23.0396%	12.0325%
---------	---------	----------	----------	----------	----------

+ info

Ver cines en que se proyecta.

Woodvie

CONFIGURACIÓN
ACERCA DE

Juegos
Slide FunSpace
Ver más -

Anuncios Crear un anuncio

I.1.1.1.4 VER INFORMACIÓN SOBRE LOS CINES

Se trata de una sección análoga a la anterior, Ver las películas en cartelera, es decir, es necesario tener seleccionada la provincia. Cuando se haya completado este paso, para consultar dicha información el usuario debe acceder al apartado “**Ver cines**” del menú principal. Una vez dentro, podrá ampliar la información sobre el cine que le interese, visionando, entre otras cosas, qué películas se están proyectando en dicho cine, o un mapa interactivo basado en “Google Maps”.

facebook Buscar Inicio Perfil Buscar amigos Cuenta

Inicio Obtener recomendación Ver cartelera Ver cines Realizar test

CINES

A continuación puedes ver todos los cines de tu provincia disponibles en nuestra base de datos.

Elige una zona de tu provincia

Alcobendas Ir

Prev 1 Next

[Ábaco Alcobendas 3D](#)

Dirección: Avenida De Bruselas, 21 , 28108, Alcobendas
Ciudad: Alcobendas
Provincia: Madrid
Ver películas
Midnight in Paris, Sin identidad, Piratas del Caribe: En mareas misteriosas, X-Men: Primera generación, Hanna, Algo prestado, Kung Fu Panda 2, Un cuento chino, Blitz, Resacón 2, ¡Ahora en Tailandia!, Sólo una noche,

[Cinesa La Moraleja 3D](#)

Dirección: Avenida Europa, 13-15 , 28109, Alcobendas
Ciudad: Alcobendas
Provincia: Madrid
Ver películas

CONFIGURACIÓN
ACERCA DE

Juegos
Slide FunSpace
Ver más

Anuncios Crear un anuncio

I.1.1.1.5 CONOCER CUÁNDO SE PROYECTA CADA PELÍCULA

Una vez más, para poder resolver esta cuestión es necesario que el usuario haya seleccionado una provincia. Desde ese momento, cada vez que al usuario se le muestre información sobre una película, tendrá un enlace “[Ver cines en que se proyecta](#)”, a través del cual podrá acceder a todos los cines de su provincia que proyecten dicha película, y las sesiones (día y hora) asociadas a ellos.

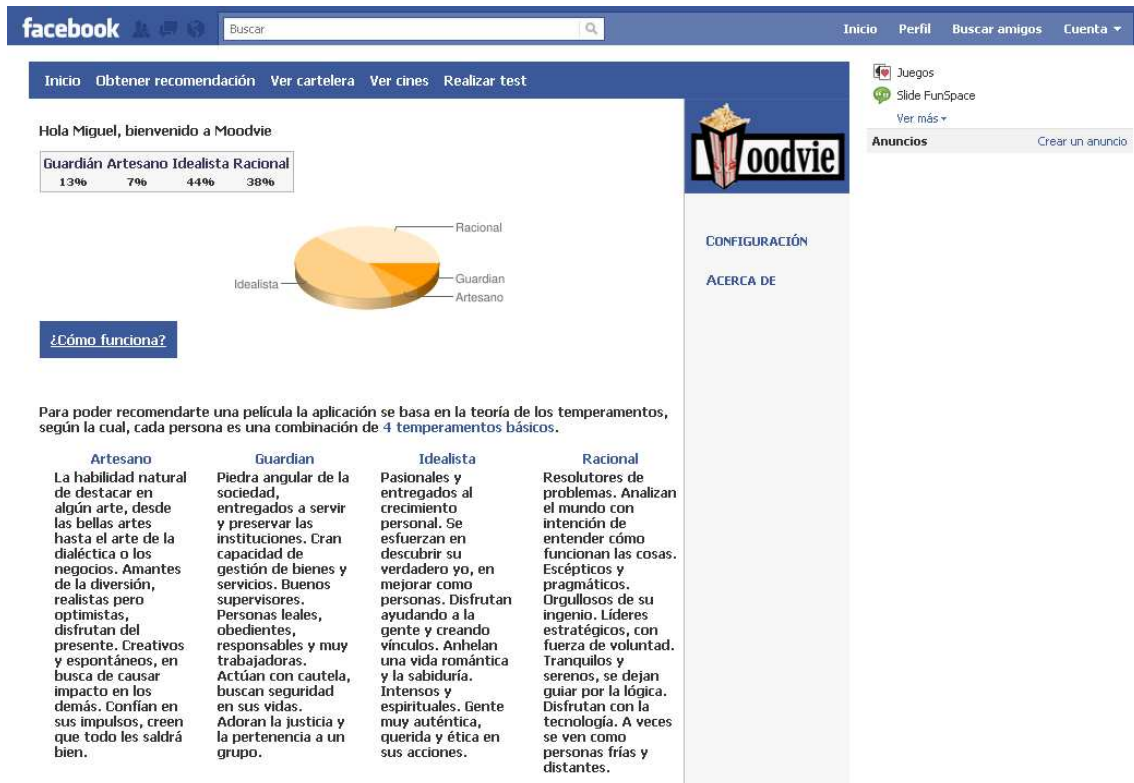
The screenshot shows the Moodvie website interface. At the top is a Facebook navigation bar with a search bar and links for Inicio, Perfil, Buscar amigos, and Cuenta. Below this is a navigation menu with links: Inicio, Obtener recomendación, Ver cartelera, Ver cines, and Realizar test. The main content area is titled "CINES" and displays information for "UGC Cinesa Méndez Álvaro". The cinema's address is "Dirección: Calle Acanto, 2, 28045, Madrid", the city is "Ciudad: Madrid", and the province is "Provincia: Madrid". A link for "Midnight in Paris" is provided. Below this is a table showing the days of the week and the corresponding cinema sessions (times).

Día	Hora
Viernes	12:00 14:00 16:00 18:05 20:15 22:30
Sábado	12:00 14:00 16:00 18:05 20:15 22:30
Domingo	12:00 14:00 16:00 18:05 20:15 22:30
Lunes	12:00 14:00 16:00 18:05 20:15 22:30
Martes	12:00 14:00 16:00 18:05 20:15 22:30
Jueves	12:00 14:00 16:00 18:05 20:15 22:30

On the right side of the page, there is a "CONFIGURACIÓN" section with a link "ACERCA DE". Above this, there is a "Juegos" section with a link "Slide FunSpace" and a "Ver más" link. Below that is an "Anuncios" section with a link "Crear un anuncio".

I.1.1.6 OBTENER UNA RECOMENDACIÓN

Para poder obtener una recomendación, basta con que el usuario haya respondido el test –es decir, tenga un temperamento asociado- y seleccionado la provincia en que se encuentra. Con dichos datos, la página de inicio del usuario será similar a la siguiente:



Desde ese momento, el usuario tan sólo debe pulsar en “Obtener recomendación” para que se le muestren dos opciones: Recomendación normal o basada en ánimo. En caso de seleccionar esta segunda, aparecerá un menú desplegable en la parte inmediatamente inferior, donde podrá seleccionar el estado de ánimo que desea. Después, tras pulsar sobre el botón “Recomiéndame”, el usuario podrá ver una lista de hasta cinco películas, pudiendo ser menor, que se le recomiendan. Es importante señalar que aparecen en estricto orden, siendo la primera película una con mayor puntuación que la segunda, y así sucesivamente. Es por ello que es posible que las últimas películas de la recomendación no se ajusten especialmente al gusto del usuario.

The screenshot shows the 'RECOMENDADOR' section of the Moodvie interface. It prompts the user to choose between 'Recomendación normal' and 'Recomendación basada en ánimo'. The second option is selected. Below this, there is a dropdown menu for '¿Qué estado de ánimo buscas?' with 'Relajado' selected. A button labeled '¡Recomiéndame!' is visible. The interface includes the same navigation links and right-hand sidebar as the previous screenshot.

RECOMENDADOR

A continuación, podrás elegir entre los dos tipos de recomendación que ofrecemos, normal y basado en ánimo.

☐ Recomendación normal

☒ Recomendación basada en ánimo

¿Qué estado de ánimo buscas? Relajado

I.1.2 Administrador

El administrador debe realizar, fundamentalmente, la tarea de actualización de la base de datos. Para ello, empleará la aplicación Moodvie – Extractor de información, en lugar de la aplicación de Facebook. En ella, básicamente se pueden realizar las siguientes tareas:

1. Crear la Base de Datos
2. Eliminar los registros de la Base de Datos
3. Consultar información de los registros
4. Actualizar películas de una/todas las provincia/s
5. Actualizar cines de una/todas las provincia/s
6. Actualizar sesiones de una/todas las provincia/s
7. Actualizar por completo una/todas las provincia/s

I.1.2.1 CREAR LA BASE DE DATOS

Para realizar esta tarea, basta con acceder a la pestaña **Administrar** de la aplicación Moodvie – Extractor de información y pulsar sobre el botón “**Crear BD**”.

I.1.2.2 ELIMINAR LOS REGISTROS DE LA BASE DE DATOS

Desde la pestaña **Administrar**, pulsar sobre el botón Eliminar entradas. Si se actualiza la información, se podrá ver entonces que todos los campos de información contendrán 0 registros.

I.1.2.3 CONSULTAR INFORMACIÓN DE LOS REGISTROS

De nuevo, en la pestaña **Administrar**, el administrador deberá pulsar sobre el botón que tiene dos flechas entrelazadas y forman un círculo. Así, la información sobre el número de películas, cines y sesiones se actualizará a las disponibles en ese momento.

I.1.2.4 ACTUALIZAR PELÍCULAS DE UNA/TODAS LAS PROVINCIA/S

En la pestaña **Actualizar**, marcar la opción **Películas**. De ese modo, se introducen en la base de datos todas las películas que se proyecten en ese momento en la provincia seleccionada. En caso de querer actualizar todas las provincias, se deberá seleccionar “**Todas**” en la lista desplegable de selección de provincia.

I.1.2.5 ACTUALIZAR CINES DE UNA/TODAS LAS PROVINCIA/S

En la pestaña **Actualizar**, marcar la opción **Cines**. De ese modo, se introducen en la base de datos todos los cines de la provincia seleccionada. En caso de querer actualizar todas las provincias, se deberá seleccionar “**Todas**” en la lista desplegable de selección de provincia.

I.1.2.6 ACTUALIZAR SESIONES DE UNA/TODAS LAS PROVINCIA/S

En la pestaña **Actualizar**, marcar la opción **Sesiones**. De ese modo, se introducen en la base de datos todas las sesiones de cada uno de los cines de la provincia seleccionada. En caso de querer actualizar todas las provincias, se deberá seleccionar “**Todas**” en la lista desplegable de selección de provincia.

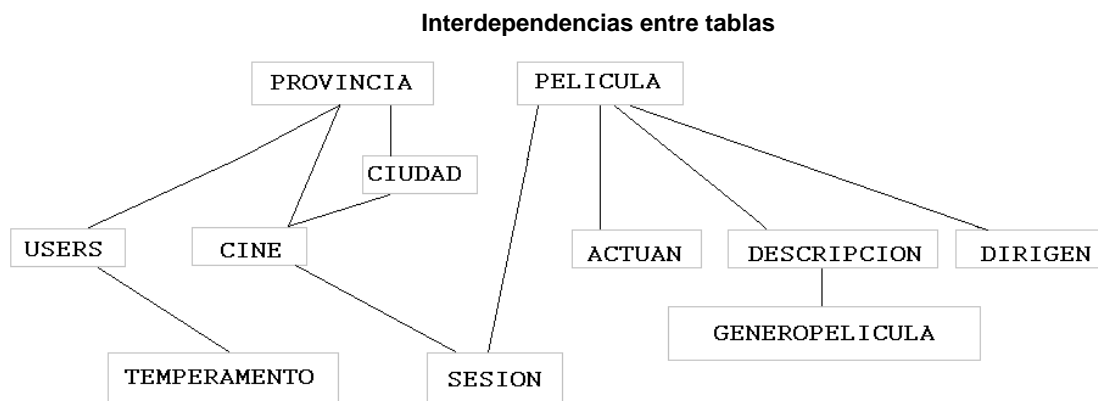
I.1.2.7 ACTUALIZAR POR COMPLETO UNA/TODAS LAS PROVINCIA/S

En la pestaña **Actualizar**, marcar la opción **Toda la Provincia**. De ese modo, se introducen en la base de datos todas las películas, cines y sesiones de la provincia seleccionada. En caso de querer actualizar todas las provincias, se deberá seleccionar **“Todas”** en la lista desplegable de selección de provincia.

I.1.2.8 INFORMACIÓN IMPORTANTE SOBRE LA ACTUALIZACIÓN

Es recomendable, dado el nivel de dependencias entre tablas por las claves foráneas (FK) utilizadas, utilizar este último método (Actualizar por completo una/todas las provincia/s) para realizar las actualizaciones, ya que no realizar las actualizaciones de forma correcta puede provocar fallos en las inserciones. Por ejemplo, tratar de actualizar los cines y sus proyecciones (o pases) antes de introducir las películas puede llevar a que dicha película no exista aún en la BBDD y por tanto se produzca una violación de las restricciones de integridad.

En el siguiente gráfico se muestran las dependencias entre tablas. Las tablas son mostradas como rectángulos grises con sus nombres escritos, y las dependencias son representadas mediante las líneas que las unen. Así, la tabla que se encuentre en la parte inferior depende de la que se encuentre en la superior.



La ejecución recomendada es la actualización de “Toda la provincia”, que realizará la actualización de la provincia en el orden adecuado –esto es, primero películas y a continuación los cines y sus sesiones-. En cualquier caso, se permite realizar la actualización únicamente de películas, de sesiones o de cines.

Sobre las provincias a actualizar, se deja a elección la actualización de una única provincia a modo individual, o de todas las provincias españolas.

Debido al volumen de datos que se manejan, las actualizaciones son costosas en tiempo. Es especialmente crítica la primera de todas, puesto que, como ya se explicó en la sección “Extracción de información”, cada vez que se actualizan las películas de una provincia, se consulta previamente si dicha película ya existe en la BBDD, y en caso afirmativo, se obvia volver a descargar toda la información asociada. En la primera actualización, dado que no existe ningún registro en la base de datos, se deben descargar todas las que se estén proyectando en la provincia. En posteriores actualizaciones se descargarán únicamente las películas de nuevo estreno, pero no las que ya se hubieran estrenado, puesto que ya estarán introducidas en la BBDD.

El coste en tiempo en cines y sesiones es similar entre la primera ejecución y las restantes, puesto que siempre se actualizan todos los datos.

1.2 Manual del sistema

En este apartado, se detallarán los requisitos para hacer funcionar las distintas aplicaciones, así como una guía de instalación.

1.2.1 Requisitos del sistema

La aplicación Moodvie – Extractor de información, para su correcto funcionamiento, necesita los siguientes componentes:

1. Windows XP Service Pack 1
2. Java JDK 1.6
3. MySQL 5.1.53 – Community

El sistema recomendador ha sido probado utilizando los siguientes componentes:

1. Windows XP Service Pack 1
2. Java JDK 1.6
3. MySQL 5.1.53 – Community
4. Xampp BasisPackage version 1.7.4
5. Apache Tomcat 7.0.12
6. JColibri 2.1
7. PHP/Java Bridge 6.21

1.2.2 Instrucciones de instalación

En esta sección, se detallarán los pasos para hacer funcionar ambas aplicaciones.

I.2.2.1 MOODVIE – EXTRACTOR DE INFORMACIÓN.

Una vez instalada la Java JDK1.6 y MySQL 5.1.53 Así mismo, es necesario que el servidor MySQL esté activo (las pruebas se han hecho sobre el puerto 3306) en el ordenador local, y que exista un usuario con permisos de administrador.

Sin embargo, otros puertos o pares de nombre_de_usuario/contraseña son aceptados por la aplicación. A este respecto, la unión de la ruta:puerto y el par nombre_de_usuario/contraseña en la aplicación es llamada ParamsConexionBD, y se puede modificar en cualquier momento. Estos parámetros son, además, necesarios para la invocación de cualquier método público de la clase BD, que es la encargada de realizar las conexiones.

Para facilitar la utilización del extractor de información se ha desarrollado un front-end gráfico que simplifica el uso, pudiendo utilizar las funciones más comunes (de actualización, consulta de registros o borrado de la base de datos).

Puesto que necesita de la base de datos, es necesario tener el servicio de MySQL ejecutándose para un funcionamiento correcto. De igual modo, debido a que las tablas que se van a acceder aún no están creadas, el primer acceso a la pestaña “Administrar” será fallido. Por ello, para poder introducir registros y actualizar la base de datos, es necesario crearla. Desde ese momento, en el que se crean todas las tablas y se rellena una de ellas, la de provincias, se podrán ver campos como la última vez que se ha actualizado la base de datos, el número de películas que en la actualidad figuran en cartelera, el número de salas de cine que se tienen y de proyecciones (sesiones) de las películas en dichas salas.

I.2.2.2 RECOMENDADOR

Para hacer funcionar el recomendador, es necesario instalar todos los componentes antes expuestos. A continuación, se debe proceder a realizar las siguientes tareas.

NOTA

En el manual, utilizaremos variables para definir las rutas del sistema, de modo que sea fácilmente intercambiable. Como se podrá observar, las rutas relativas que se ofrecen en esta guía son pertenecientes a entornos Windows. El proceso de instalación en Linux es similar, con las peculiaridades de este sistema.

\$extractor= C:/Workspaces/Workspace Java/SSII

\$xampp= C:/xampp

\$tomcat=C:/apache-tomcat-7.0.12

1. Editar el archivo **\$xampp**/apache/conf/httpd.conf, y dentro del apartado `<IfModule alias_module>` añadir lo siguiente:

```
Alias /carteles "$extractor/carteles"
<Directory "$extractor/carteles">
    Options Indexes FollowSymLinks Includes ExecCGI
    AllowOverride All
    Order allow,deny
    Allow from all
</Directory>
```

2. Editar el archivo **\$xampp**/php/php.ini, y descomentar la siguiente línea:

```
extension=php_curl.dll
```

3. Lanzar Apache Web Server

4. Introducir la carpeta de la aplicación Java recomendadora (recomendador) en la siguiente ruta:

```
$extractor/bin;
```

5. Editar el archivo **\$tomcat**/conf/tomcat-users.xml. En él, añadir las siguientes líneas dentro de los tagtomcat-user:

```
<role rolename="tomcat"/>
<role rolename="manager"/>
<user username="usuario" password="pass" roles="tomcat,manager"/>
```

6. Editar el archivo `$tomcat/bin/catalina.bat`, y añadir tras el texto

```
rem Ensure that any user defined CLASSPATH variables are not used on startup,
rem but allow them to be specified in setenv.bat, in rare case when it is needed.
```

la siguiente línea:

```
set "CLASSPATH=%CLASSPATH%;"

Si la configuración se estuviese realizando para una distribución de Linux,
entonces el archivo a modificar sería $tomcat/bin/catalina.sh, y el lugar en el que
añadir la ruta sería el siguiente al texto;

# Ensure that any user defined CLASSPATH variables are not used on startup,
# but allow them to be specified in setenv.sh, in rare case when it is needed.

CLASSPATH= $CLASSPATH
```

7. Crear una nueva variable de sesión, con nombre `CLASSPATH`. En ella, introducir las siguientes rutas:

```
$extractor\bin; $extractor\SSII\Librerias\mysql-connector-java-5.1.13-bin.jar
```

8. Lanzar apache-tomcat 7.0.12. Para ello, en Windows, basta con ejecutar `$tomcat/bin/startup.bat`. En Linux, `$tomcat/bin/startup.sh`

9. Desde el navegador web, acceder a la dirección `localhost:8080`, es decir, al panel de administración de Apache-Tomcat. Allí, identificarse con el usuario/pass antes indicado (punto 5 de esta guía de instalación) dentro de la sección Administration/Tomcat Manager.

10. Realizar un despliegado de PHP/Java Bridge 621.war. Para ello, dentro del apartado desplegar (Deploy), introducir lo siguiente:

Trayectoria de Contexto (opcional): `/JavaBridge`

URL de archivo de Configuración XML:

URL de WAR o Directorio: Ruta donde se encuentre PHP/Java Bridge 621.war

11. Ahora, debería haber un nuevo apartado, con nombre `/JavaBridge`, dentro del manager de Tomcat.

12. Llevar la carpeta con el proyecto web (moodvie) a la siguiente ruta:

```
$xampp\htdocs
```

13. Desde este momento, la aplicación web debería estar funcionando correctamente.

NOTA

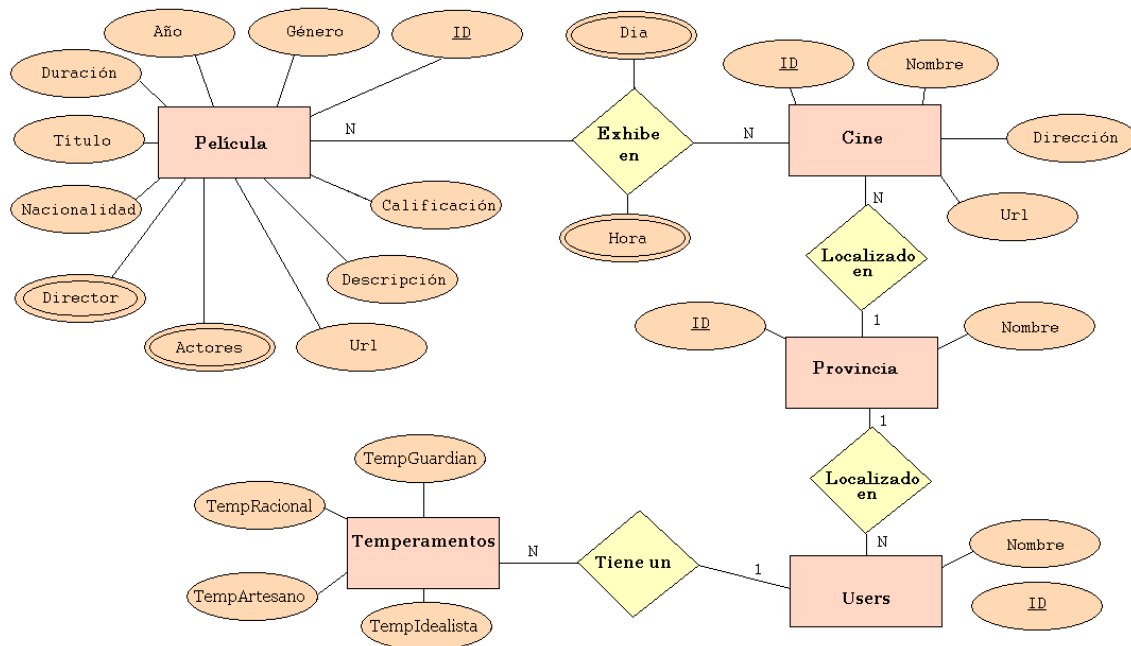
Además de todos los pasos arriba comentados, es necesario tener creado un espacio para la aplicación en Facebook. Para más información, consulte el apartado 6.2 Registro en Facebook.

ANEXO II. DISEÑO DE LA BASE DE DATOS

En esta sección se detalla el diseño de la base de datos que utiliza la aplicación.

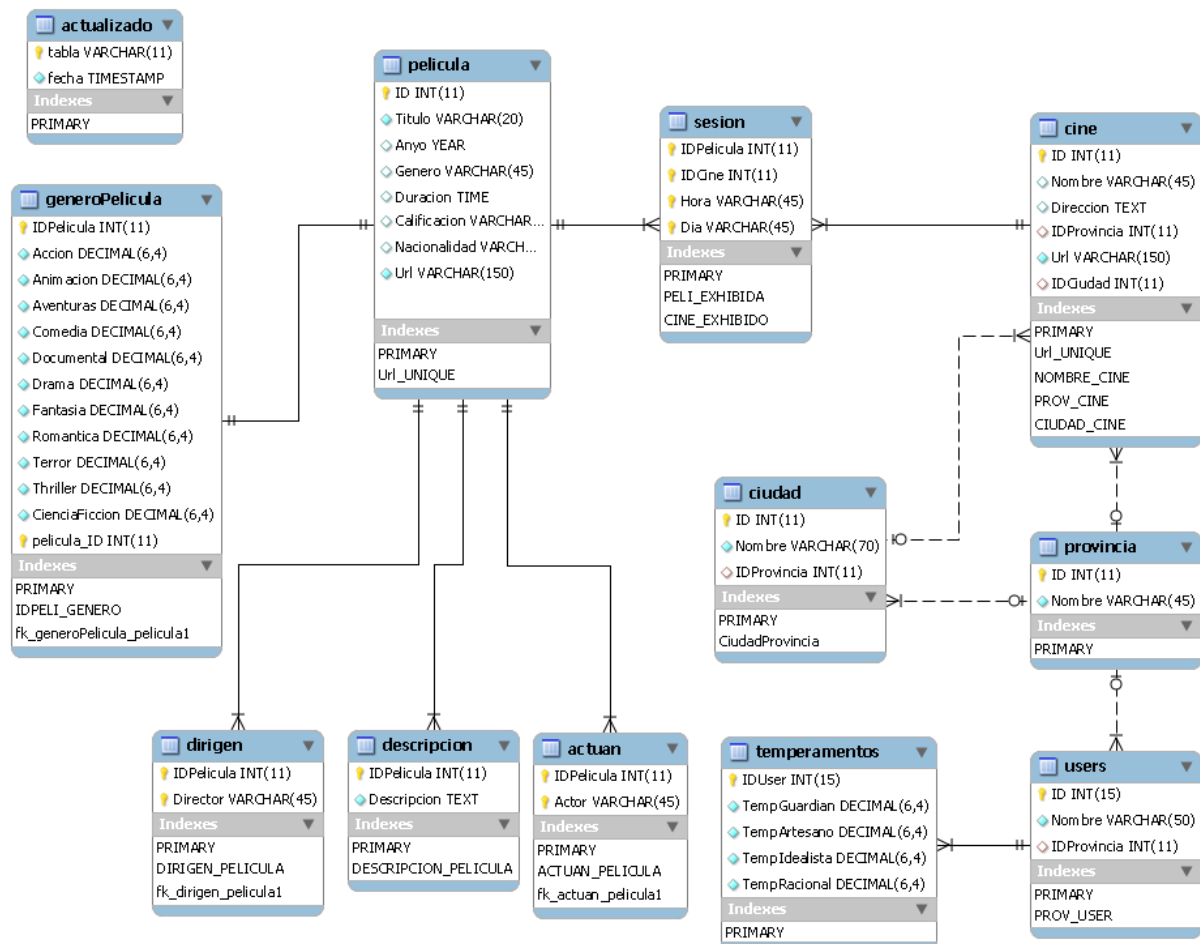
II.1 Modelo Entidad Relación

El modelo Entidad-Relación de la base de datos es el siguiente.



II.2 Modelo Relacional

Pelicula	(<u>ID</u> , <u>Anyo</u> , <u>Titulo</u> , <u>Duracion</u> , <u>Calificacion</u> , <u>Nacionalidad</u> , <u>Genero</u> , <u>Url</u> , <u>Rutalmg</u>)
Sesion	(<u>IDPelicula</u> , <u>IDCine</u> , <u>Hora</u> , <u>Dia</u>)
Cine	(<u>ID</u> , <u>Nombre</u> , <u>Direccion</u> , <u>IDProvincia</u> , <u>Url</u> , <u>IDCiudad</u>)
Ciudad	(<u>ID</u> , <u>Nombre</u> , <u>IDProvincia</u>)
Provincia	(<u>ID</u> , <u>Nombre</u>)
Actuan	(<u>IDPelicula</u> , <u>Actor</u>)
Dirigen	(<u>IDPelicula</u> , <u>Director</u>)
Descripcion	(<u>IDPelicula</u> , <u>Descripcion</u>)
Users	(<u>ID</u> , <u>IDFB</u> , <u>Nombre</u> , <u>IDProvincia</u>)
Temperamentos	(<u>IDUser</u> , <u>TempGuardian</u> , <u>TempArtesano</u> , <u>TempIdealista</u> , <u>TempRacional</u>)
Actualizado	(<u>tabla</u> , <u>fecha</u>)
GeneroPelicula	(<u>IDPelicula</u> , <u>Accion</u> , <u>Animacion</u> , <u>Aventuras</u> , <u>Comedia</u> , <u>Documental</u> , <u>Drama</u> , <u>Fantasia</u> , <u>Romantica</u> , <u>Terror</u> , <u>Thriller</u> , <u>CienciaFiccion</u>)



Cabe destacar que se ha introducido la tabla descripción para aumentar la eficiencia en las consultas, pues se tratará de un campo con mucho texto y no consultado tan a menudo como otros campos.

El criterio elegido a la hora de nombrar los campos, especialmente aquellos que son claves foráneas, es el de respetar el nombre de la clave referenciada y concatenarle el nombre de la tabla que la contiene.

Así mismo, para mejorar la eficiencia y eliminar una tabla, en la tabla *Cine* se introduce una referencia a la provincia.

No parece útil, debido el enfoque de las recomendaciones, crear una tabla de Actores o de Directores, a pesar de que estos puedan repetirse.

Se ha introducido el atributo Url tanto en Película como en Cine. Esto es así porque facilitan enormemente la introducción y actualización de la base de datos, puesto que tienen un índice de tipo “unique”, y así se evitan repeticiones. Este atributo Url hace referencia a la dirección de “La Guía del Ocio” de donde han sido obtenidos los datos.

II.2.1 Índices

Tabla Película:	<p>Índice primario sobre ID. (PK)</p> <p>Índice TITULO_PELI sobre Título. (Árbol B)</p> <p>Índice unique Url_UNIQUE sobre Url.</p>
Tabla Sesión:	Índice primario sobre las 4 columnas (PK)
Tabla Cine:	<p>Índice primario sobre ID. (PK)</p> <p>Índice NOMBRE_CINE sobre Nombre. (Árbol B)</p> <p>Índice unique Url_UNIQUE sobre Url.</p>
Tabla Provincia:	Índice primario sobre ID (PK)
Tabla Actuan:	Índice primario sobre las 2 columnas (PK)
Tabla Dirigen:	Índice primario sobre las 2 columnas (PK)
Tabla Descripción:	Índice primario sobre IDPelícula (PK)
Tabla Ciudad:	Índice primario sobre ID (PK)
Tabla Users:	<p>Índice primario sobre ID (PK)</p> <p>Índice unique IDFB_UNIQUE sobre IDFBI.</p>
Tabla Temperamentos:	Índice primario sobre IDUser (PK)
Tabla Actualizado:	Índice primario sobre Tabla (PK)
Tabla GeneroPelícula:	Índice primario sobre IDPelícula (PK)

Además, por cada Foreign Key se crea un índice con el mismo nombre, que agiliza la consulta y actualizado de restricciones de integridad.

II.2.2 Foreign Keys

Tabla Película:	No tiene.
Tabla Sesión:	<p><i>PELI_EXHIBIDA</i> relaciona <i>IDPelícula</i> con <i>Película.ID</i> On Update, Cascade. On Delete, Restrict</p> <p><i>CINE_EXHIBIDO</i> relaciona <i>IDCine</i> con <i>Cine.ID</i> On Update, Cascade. On Delete, Restrict</p>
Tabla Cine:	<i>PROV_CINE</i> relaciona <i>IDProvincia</i> con <i>Provincia.ID</i> . On Update, Cascade. On Delete, Restrict
Tabla Provincia:	No tiene.
Tabla Actuan:	<i>ACTUAN_PELICULA</i> relaciona <i>IDPelícula</i> con <i>Película.ID</i> On Update, Cascade. On Delete, Cascade.
Tabla Dirigen:	<i>DIRIGEN_PELICULA</i> relaciona <i>IDPelícula</i> con <i>Película.ID</i> On Update, Cascade. On Delete, Cascade
Tabla Descripción:	<i>DESCRIPCION_PELICULA</i> relaciona <i>IDPelícula</i> con <i>Película.ID</i> On Update, Cascade. On Delete, Cascade
Tabla Ciudad:	<i>CiudadProvincia:</i> relaciona <i>IDProvincia</i> con <i>Provincia.ID</i> On Update, Cascade. On Delete, Cascade
Tabla Users:	<i>PROV_USER</i> relaciona <i>IDProvincia</i> con <i>Provincia.ID</i> On Update, Cascade. On Delete, Cascade.
Tabla Temperamentos:	<i>TEMP_USER</i> relaciona <i>IDUser</i> con <i>Users.ID</i> On Update, Cascade. On Delete, Cascade.
Tabla Actualizado	No tiene.
Tabla GeneroPelícula:	<i>PELI_GENERO</i> relaciona <i>IDPelícula</i> con <i>Película.ID</i> On Update, Cascade. On Delete, Cascade.

